A Comparative Study of BRISK, ORB and DAISY Features for Breast Cancer Classification

Ghada Ouddai, Ines Hamdi and Henda Ben Ghezala
RIADI Laboratory, National School of Computer Science (ENSI), University of La Manouba, La Manouba, Tunisia

Keywords: Histopathological Image Processing, Feature Extraction, Binary Robust Invariant Scalable (BRISK), Oriented FAST and Rotated BRIEF (ORB), DAISY Descriptor, Bag-of-Features (BoF), Machine Learning.

Abstract: Medical data analysis is one of the most emergent fields over the past decades. In Digital histopathology, images are analysed, mainly, to detect disease or tumors and identify their types and grade. One of the most used practices in this field is the feature extraction. In this paper, we propose the application of BRISK, ORB and BRISK/DAISY on RGB histological images. The purpose of this work is to recognise the breast tumor type (benign or malignant). These features extractors are combined with BoF by kmeans and SVM. A limited amount of images is used during the training of the system. Out of the three methods, Color-BRISK/BoF/SVM solution gave the best accuracy value (72.5%) while Color-ORB/BoF/SVM was the fastest.

1 INTRODUCTION

Histopathology, which is also known as pathological histology, is a bio-medical field that offers a useful techniques for cancer and disease detection and grading. Histology and histopathology share the same sample preparation, called histological process, and the same study tool: the microscope. The difference between these two sub-fields is the purpose of study: in histology, samples as analyzed to observe the cells morphological development, on the other side, a sample study in histopathology is performed to detect abnormal tissues and diseases. When examined for pathological purposes, a histological sample can be very effective for detection of tumors as well as defining its nature and grade.

Following the digitalization of medical data, an emergence of AI tools applications to these latter has been observed. In digital histopathology, the main domain data is the histopathological image, which is generated by scanning a given specimen. Depending on the tools, stains and staining techniques used during the histological process, the image processing method is selected. In fact, in histology, their are various tools for sample cutting, preparation and staining. The three known-to-date staining techniques are: histochemistry (HC), immunohistochemistry (IHC) and immunofluorescence (IF), for each one there are hundreds of possible stains. The choice of stains depends on the target cell and study context; each stain or stains combination allows the emphasis and highlight of certain morphological parts, the frequently used ones are Eosin (E), Hematoxylin (H) and their combination (H&E).

Over the past years, there was a huge number of researches and attempts to create the perfect autonomous Computer-Assisted Diagnosis (CAD) system for disease and cancer detection and grading using histopathological images. In parallel to that, numerous works focused on the images retrieval and/or registration. Our study of the state-of-the-art works in histopathological image analysis field, such as presented in the papers (Azevedo Tosta et al., 2017), (Das et al., 2020), (Gurcan et al., 2009), (Irshad et al., 2014), (Komura and Ishikawa, 2018), (Li et al., 2020), (Ai et al., 2021), showed that the majority of CADs proposed in this field depends on deep learning methods. These latter offer great classification results however there are some limitations to them:

- First, to achieve good result, a considerable number of labelled data should be used. The more images is analysed, the good performance is obtained.
- Second, deep learning methods need powerful computation machines. GPU-based calculation offers fast CNN and RNN training however, if the experiments data-set is very large, a memory overflow can occur. In other hand, CPU-based calculation is slow but the memory is unlimited.

Ouddai, G., Hamdi, I. and Ben Ghezala, H.
A Comparative Study of BRISK, ORB and DAISY Features for Breast Cancer Classification.
DOI: 10.5220/0011902200003411
In *Proceedings of the 12th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2023)*, pages 964-970
ISBN: 978-989-758-626-2; ISSN: 2184-4313

The available labelled histopathological data-sets are mainly consisted of HC H&E slides. There are some IHC data-sets but, to the best of our knowledge, there is no dedicated IF images collections. The majority of these image-bases regroup breast tumor slides either by type (benign, malignant) or by grade. In this work, to classify breast tumoral cells, we use and compare three features extractors: Binary Robust Invariant Scalable (BRISK) (Leutenegger et al., 2011), Oriented FAST and Rotated BRIEF (ORB) (El-Hallak and Lovell, 2013) and BRISK-keypoints/DAISY-descriptors. These methods are applied to RGB images rather than gray-scale images; this allows the exploitation of color information. For each feature extractor, an encoding on Bag-of-Features (BoF) by kmeans/frequency histogram is performed before the last step of classification by SVM (Support Vector Machine). A CPU-based calculation and a limited data-set are used in the experiments. The main purpose of this work is to find which feature extractor is the fastest and the more accurate in identifying the tumor type.

This paper is organized as follows: in section 2, we give a quick overview of literature works and features extraction applications in digital histopathology, then, in section 3, we introduce and explain our approach in details beginning by the pre-processing method used till the classification system. The proposed approach is evaluated in section 4 where all the specifications of computation architecture, data-set and test results are listed. In section 5, we give a conclusion of this paper and a perspective of future works.

2 BACKGROUND

2.1 Low-Level Information Detection

In digital histopathology, image processing techniques are used mainly for segmentation or Regions-of-Interest (ROI) detection. Features extraction is less used. In (Öztürk and Akdemir, 2018), the authors study the efficiency of the combination of different feature extractors with a variety of classifier. For the texture characteristics, Gray-Level Co-occurrence Matrix (GLCM), Gray-Level Run Length Matrix (GLRLM) and Segmentation-based Fractal Texture Analysis (SFTA) are calculated. For the luminance features, the authors used Local Binary Pattern (LBP). The Local Binary Gray Level Co-occurrence Matrix (LBGLCM) is also calculated for common texture/luminance features. The evaluation of these methods was performed by the authors using some common classifiers such as Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Linear Discriminant Analysis (LDA) and Boosted Trees. The best performance was achieved by the SFTA/Boosted Trees system.

Local Binary Pattern (LBP) was used in (Kumar et al., 2018) combined with Bag-of-Visual-Words (BoVW) and SVM; a comparison with LBP deep features was established. A variant of LBP, named mrcLBP which consists of calculating LBP on each RGB channel separately, was used in (Simon et al., 2018). Other than that, KAZE features were used in (Sanchez-Morillo et al., 2018) to classify breast cancer H&E-stained images. In (Popovici et al., 2016), the authors propose to use directly the clustering by kmeans to construct a local Bag-of-Features (BoF) from the image, named code blocks. these latter are jointed to tumor size, grade and gene expression.

Scale-Invariant Feature Transform (SIFT) (Lowe, 1999) was used in (Li et al., 2019b) and (Li et al., 2019a) alongside other methods to prepare cervical histopathological images for classification. In (Irshad et al., 2013), to detect mitosis from H&E stained images, the authors proposed an system based on SIFT and texture features to detect the key-points from R and B channels. In (Bukała et al., 2020), the authors proposed the exploitation of color information by using and comparing various Color-SIFT. Similar procedure was performed in (Ouddai et al., 2023) where the authors used RGB-SIFT to classify breast cancer using a small data-set.

2.2 Databases

For CAD, databases are needed to train the machine/deep learning systems. Digital histopathological databases regroup similar images: the study context, size (Whole slide image WSI, patches or regular sized images) and stains used (histochemistry (HC), immunohistochemistry (IHC), immunofluorescence (IF), Eosin (E), Hematoxylin (H) or their combination (H&E)) must be the same. Databases in general need to be labeled by field experts. In the case of digital histopathology, the size of image, stains used and studied cells must be provided with the images, some additional details such as: age, gender, health situation... etc. can be useful in some studies. In Tab. 1, we present some of the existing histopathological databases.

Table 1: Examples of Histopathological Image Databases.

Name	Type	Stains used	Cell	Dataset size
BreCaHAD (Aksac et al., 2019)	Regular images	H&E	Breast	162
(Wang et al., 2022)	WSI	H&E	ovarian Cancer	288
BreakHis (Spanhol et al., 2016)	Regular image	H&E	Breast	7909 images
Medisp HICL (Glotsos et al., 2008), (Kostopoulos et al.,), (Ninos et al., 2016)	Regular image	H&E/IHC	Brain, Breast Larynx	3870 images
Camelyon17 (Litjens et al., 2018)	WSI	H&E	Breast	1339 images
KIMIA Path24 (Shafiei et al., 2021)	WSI patches	H&E/IHC	/	22591 (train patches) 1325 (test patches)

3 METHODOLOGY

As stated before, in digital histopathology, color information is very important; it indicates the histological process staining results. The choice of staining techniques and stains depends on the cells or disease to detect and study, in fact, for each case exists one or multiple adequate pigmentation method. The choice of this latter is an important step in the histological process.

In this paper, we propose the modification of regular features descriptors (in our case: BRISK, ORB and DAISY) to extract key-points from RGB channel rather than gray-scale single channeled image. Alongside that, we use an encoding method to regroup descriptor into clusters of same nature. In the final step, a supervised classification method is used to determinate and interpret the input image nature. The context of our study is the classification of breast tumor by type (benign/malignant). The main purpose of this work is to present answers to the following research questions:

- Between color-BRISK descriptor, color-ORB descriptor and color-BRISK-keypoints/DAISY-descriptor, which method is the more appropriate to histopathological images?
- When dealing with a limited database, which method can offer a better slide classification scores?
- For each method and for a CPU-based calculation, what is the maximal execution time to be expected?

3.1 Pre-Processing

In any computer vision sub-domain, the preparation of input image for further processing, analysis and interpretation is really important; the pre-processing tools and method must be chosen thoroughly. In digital histopathology in particular, the nature of image is delicate due to the morphological textures of cells and tissues. The most frequent noise that can occur during the scanning of histological slides are green hues or shadows and luminance unbalance.

The database selected in our study consists mainly of HC H&E stained slides. When analysing these samples, the first global remark is that the majority of images contain a green shadow; its intensity differs from image to another. To remedy to this problem and eliminating the hue without loosing important morphological textures, we use a lightweight pre-processing method; we chose the bias and gain function (see Eq. 1) to correct the luminance and contrast.

$$Output(i,j) = \alpha * Input(i,j) + \beta \qquad (1)$$

In the definition of the equation above, the parameter α and β are fixed in an experimental way: α controls the contrast while β controls the brightness. For the parameter α, its value must be between 1 and 3; if $\alpha < 0$, the result image colors will be compressed. For the parameter β, its value should range between 0 and 100. The procedure of selection of these latter's values depends on the nature of the studied images. In the case of HC H&E stained slides, we noticed that a green shadow appears more often than in IHC images. Our first tentative of adjusting the images quality using the same parameter's values on the three channels was unsatisfactory. This latter led to the fading of some important morphological and color details.

After numerous experiments, we found that adjusting each channel of the RGB image separately is the best solution to obtain well calibrated contrast and brightness. For the HC H&E slides, we fix $\alpha = 1.2$ and $\beta = 25$ for the R and B channels, for the G channel, we fix $\alpha = 1.2$ and $\beta = -25$. In the case of IHC image, we fix $\alpha = 1.1$ and $\beta = 10$, for the three channels. The pre-processing results are shown in Fig. 1.

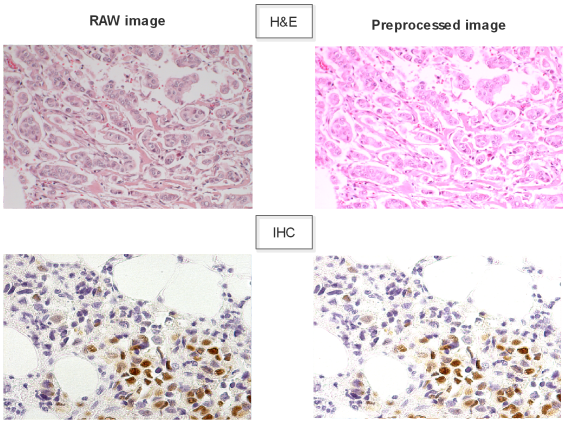

Figure 1: Results of our pre-processing methods – Examples of H&E and IHC images of Medisp HICL database ((Glotsos et al., 2008), (Kostopoulos et al.,) and (Ninos et al., 2016)).

3.2 Features Extraction

The main purpose of this work, as stated before, is to apply, evaluate and compare three methods for features extraction. These latter are slightly modified to operate on RGB image instead of gray-scale images. With this, we ensure the use of the histological staining information as well as the morphological textures. In our work, we chose the method: Binary Robust Invariant Scalable (BRISK) (Leutenegger et al., 2011), Oriented FAST and Rotated BRIEF (ORB) and a combination of BRISK key-points and DAISY descriptors.

3.2.1 Features Extraction by Color-BRISK

The original Binary Robust Invariant Scalable (BRISK), as presented in (Leutenegger et al., 2011) is applied to the gray-scale image. This method allows the detection of local key-points and their descriptors construction. The results are rotation, scaling and translation invariant. BRISK follows the same strategy as SIFT (Lowe, 1999) while being faster.

Gray-scale image is the result generated from a direct transformation of the original image and combination of its three RGB channels. Some loss of color information can occur following this conversion. To make the most use of such details, we propose a modification to the original BRISK:

- For each RGB channel, apply BRISK and extract the key-points.
- Generate the descriptor vector by BRISK for each channel.
- Concatenate the three descriptor vectors into one vector.

By this methodology, each channel of the RGB image is considerate by BRISK as a gray-scale image. In the end, and by combining the three descriptors, we are sure to conserve the color information of each channel. The final vector, resulting from the concatenation of the three separate vectors, is the new descriptive representation of the input image.

3.2.2 Features Extraction by Color-ORB

Oriented FAST and Rotated BRIEF (ORB) (El-Hallak and Lovell, 2013) is a novel approach based on the original methods Features from Accelerated Segment Test (FAST) (Rosten et al., 2010) and Binary Robust Independent Elementary Features (BRIEF) (Calonder et al., 2010). ORB, as the original works of FAST, BRIEF and BRISK, is applied to gray-scale image by calculating the FAST key-points then the BRIEF descriptors. In this work, we use RGB image channels separately to compute ORB. The procedure is the same as Color-BRISK explained above.

3.2.3 Color-BRISK Features and DAISY Descriptors

In this section, we re-use the Color-BRISK key-points. These latter are passed as inputs to the Fast Local Descriptor for Dense Matching (DAISY) (Tola et al., 2010). We chose this combination to verify if better results can be obtained by using DAISY descriptor, which is known to be fast and efficient for Bag-of-Features construction.

3.3 Features Vectors Encoding

In computer vision, descriptors vectors are effective for image matching, image retrieval or object detection. For our system, we want to use the descriptors for classification purposes: rather than classifying directly the image, low-level features are used instead as its new representations. Raw descriptor vectors can not be passed directly to the classification module; an encoding on Bag-of-Features (BoF) is necessary.

For descriptor encoding on Bag-of-Features (BoF), we use the method based on kmeans and the frequencies histogram. This method is proved to be efficient in the whole image classification task. The BoFs by kmeans/frequency histogram is performed as follows:

- After choosing the number of clusters (in our case, k=5), centroids of each cluster are randomly initialized by element of the descriptor vectors space.
- For the rest of the descriptor vectors, assign a cluster and recalculate the centroid of the cluster.

- In the end, the descriptor vector are regrouped into clusters (in our case, 5 groups)
- For the features dictionary (clusters), a histogram of frequency is assigned. The latter represents the apparition number of a given descriptor of the cluster.

3.4 BoF Classification Using SVM

First introduced in (Cortes and Vapnik, 1995), Support Vector Machine (SVM) is a widely used supervised classification and regression method. This method proposed two major contributions to the supervised data classification field. The first aspect of originality provided by SVM lies in the insurance of data separability. In fact, the authors state that if in the original definition space, the data is interleaved or overlapped, a passage to higher dimensional space secures the obtaining of a linearly separable redefinitions of the original data. The second goal of SVM is to find the optimal linear hyperplane which ensures the classification of the data while optimizing the margins. In our system, we use the classic binary SVM. This latter is applied when the experiments data is contained in twos labelled classes.

4 TESTS AND EVALUATION

4.1 Experimental Setup

The elaboration of our method was performed using a machine configured as follows: Intel® Core™ I9 10th Gen up to 5.30GHz CPU, 32 GB of RAM, NVIDIA® GeForce® GTX 2080 SUPER GPU, 512 GB SSD. As a software base, we used the Python 3.8 programming language, the library OpenCV 4.4.0 and its extra modules for the image pre-processing, features extracting and their encoding on Bag-of-Features. The training and evaluation of SVM was fulfilled using TensorFlow CPU 2.4.1.

4.2 Experiments Data

As mentioned before, in the histological process, there are three possible staining techniques: histochemistry (HC), immunohistochemistry (IHC) and immunofluorescence (IF), to each method, a multitude of stains can be associated. The most frequently used ones are: Eosin (E), Hematoxylin (H) and the combination Eosin-Hematoxylin (H&E). To the better of our knowledge, in digital histopathology, there are no IF databases and a very few IHC databases; the majority of available data-sets concern HC H&E breast tumoral slides. Due to this lack of IHC and IF image bases, we decided to use H&E stained slides in our experiments. The data-set used is BreakHis (Spanhol et al., 2016), it offers a collection of H&E stained SOB slides. The images are categorized, firstly, following magnification and then following the breast tumor type (benign or malignant). In each category, images are grouped following the cells. (See Tab. 2 for data-set details and total image number for each category)

Table 2: BreakHis Dataset Details.

Type	cell	X40	X100	X200	X400
Benign	A	114	113	111	106
	F	253	260	264	237
	PT	109	121	108	115
	TA	149	150	140	130
Malignant	DC	864	903	896	788
	LC	156	170	163	137
	MC	205	222	196	169
	PC	145	142	135	138

The cells categories of the benign class are: Adenosis (A), Fibroadenoma (F), Phyllodes Tumor (PT) and Tubular Adenoma (TA). For the malignant cells, there categories are: Ductal Carcinoma (DC), Lobular Carcinoma (LC), Mucinous Carcinoma (MC) and Papillary Carcinoma (PC). The total image contained in the database is equal to 7909 images, 2480 for benign tumor and 5429 for malignant tumor. The images size is *700 x 420 pixels*.

In our work, the image are resized to *201 x 150 pixels*. For the system training step, a total of 380 randomly selected images is used, 180 for the benign class and 180 for the malignant class. In the validation step, 120 image were used, 60 for benign and 60 for malignant; these latter were selected randomly from the original data-set. The purpose of this is to limit the training data and observe which method obtains better results.

4.3 Classification Results

The interpretation of the extracted features must be given by a trained machine/deep learning system. As state before, in this paper, we use Support Vector Machine to classify our Bag-of-Features (BoF). We evaluate our system using two criteria: classification accuracy and computation time. The results are shown in Tab. 3.

Table 3: Experiments results.

Model	Classification accuracy	Precision	Recall	Computation time
RGB-BRISK/BoF/SVM	72.5%	75.47%	66.67%	4 hours 33 minutes
RGB-ORB/BoF/SVM	65%	68.75%	55%	2 hours 50 minutes
RGB-BRISK/DAISY/BoF/SVM	57.5%	57.89%	55%	4 hours 54 minutes

4.4 Results Discussion

The results of our experiments, as shown in Tab. 3, prove that even if the learning data is limited, an average and acceptable classification accuracy can be achieved. From the table, we can retain the following:

- The best accuracy value of 72.5% was obtained by the RGB-BRISK/BoF/SVM system. This system remains, however, slightly slow.
- The highest value of precision and recall were also achieved by the RGB-BRISK/BoF/SVM system.
- The fastest system is the one based on RGB-ORB, however, this latter gave a lower classification performances compared to the RGB-BRISK one.
- The RGB-BRISK key-points/DAISY descriptors method was the slowest and achieved poor classification accuracy, precision and recall.

These results were obtained using BoF by kmeans where K=5, however, we believe that the accuracy values can increase for a greater k value (10, 20 or 50 and more). In the case of a bigger k value, it is to be expected that the computation time will drastically increase. Also, compared to deep learning architectures, such as ResNet, the RGB-BRISK/BoF/SVM system remains faster in CPU-based computation. In the literature works, such as in (Ouddai et al., 2023), ResNet18 training on similar amount of data took more than 7 hours. In the case of GPU-based computation, CNN and RNN architectures can be trained in a significantly shorter time.

5 CONCLUSION AND FUTURE WORKS

In this work, we applied different features extraction methods for breast tumoral histological slides classification. The methods used are: BRISK, ORB and BRISK/DAISY. An encoding on BoF by kmeans was performed and the classification was done by SVM. We proposed the exploitation of color information by computing features from each RGB channel separately then fusing the three in one. The obtained results showed that Color-BRISK gave the best classification accuracy, Color-ORB was the fastest and achieved an accuracy of 65% and the combination Color-BRISK/DAISY gave the worst results in both computation time and classification accuracy. For future works, we intend to exploit other hybrid features extractors, other than BRISK/DAISY. Another perspective is to apply BRISK or ORB to other color-spaces.

REFERENCES

Ai, S., Li, C., Li, X., Jiang, T., Grzegorzek, M., Sun, C., Rahaman, M. M., Zhang, J., Yao, Y., and Li, H. (2021). A State-of-the-Art Review for Gastric Histopathology Image Analysis Approaches and Future Development. *BioMed Research International*, 2021.

Aksac, A., Demetrick, D. J., Ozyer, T., and Alhajj, R. (2019). BreCaHAD: A dataset for breast cancer histopathological annotation and diagnosis. *BMC Research Notes*, 12(1):10–12.

Azevedo Tosta, T. A., Neves, L. A., and do Nascimento, M. Z. (2017). Segmentation methods of H&E-stained histological images of lymphoma: A review. *Informatics in Medicine Unlocked*, 9(May):35–43.

Bukała, A., Cyganek, B., Koziarski, M., Kwolek, B., Olborski, B., Antosz, Z., Swadźba, J., and Sitkowski, P. (2020). Classification of histopathological images using scale-invariant feature transform. *VISIGRAPP 2020 - Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 5(Visigrapp 2020):506–512.

Calonder, M., Lepetit, V., Strecha, C., and Fua, P. (2010). BRIEF: Binary robust independent elementary features. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6314 LNCS(PART 4):778–792.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.

Das, A., Nair, M. S., and Peter, S. D. (2020). *Computer-Aided Histopathological Image Analysis Techniques for Automated Nuclear Atypia Scoring of Breast Cancer: a Review*. Journal of Digital Imaging.

El-Hallak, M. and Lovell, D. (2013). ORB an efficient. *Arthritis and Rheumatism*, 65(10):2736.

Glotsos, D., Kalatzis, I., Spyridonos, P., Kostopoulos, S., Daskalakis, A., Athanasiadis, E., Ravazoula, P., Nikiforidis, G., and Cavouras, D. (2008). Improving accuracy in astrocytomas grading by integrating a robust least squares mapping driven support vector machine

classifier into a two level grade classification scheme. *Computer Methods and Programs in Biomedicine*, 90(3):251–261.

Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., and Yener, B. (2009). Histopathological Image Analysis: A Review. *IEEE Reviews in Biomedical Engineering*, 2:147–171.

Irshad, H., Jalali, S., Roux, L., Racoceanu, D., Naour, G., Hwee, L., and Capron, F. (2013). Automated mitosis detection using texture, SIFT features and HMAX biologically inspired approach. *Journal of Pathology Informatics*, 4(2):12.

Irshad, H., Veillard, A., Roux, L., and Racoceanu, D. (2014). Methods for nuclei detection, segmentation, and classification in digital histopathology: A review-current status and future potential. *IEEE Reviews in Biomedical Engineering*, 7:97–114.

Komura, D. and Ishikawa, S. (2018). Machine Learning Methods for Histopathological Image Analysis. *Computational and Structural Biotechnology Journal*, 16:34–42.

Kostopoulos, S., Glotsos, D., Cavouras, D., Daskalakis, A., Kalatzis, I., Georgiadis, P., Bougioukos, P., Ravazoula, P., and Nikiforidis, G. ANALYTICAL AND QUANTITATIVE CYTOLOGY AND HISTOLOGY ® ARTICLES Computer-Based Association of the Texture of Expressed Estrogen Receptor Nuclei with Histologic Grade Using Immunohistochemically-Stained Breast Carcinomas. Technical report.

Kumar, M. D., Babaie, M., Zhu, S., Kalra, S., and Tizhoosh, H. R. (2018). A comparative study of CNN, BoVW and LBP for classification of histopathological images. *2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017 - Proceedings*, 2018-Janua:1–7.

Leutenegger, S., Chli, M., and Siegwart, R. Y. (2011). BRISK: Binary Robust invariant scalable keypoints. *Proceedings of the IEEE International Conference on Computer Vision*, pages 2548–2555.

Li, C., Chen, H., Li, X., Xu, N., Hu, Z., Xue, D., Qi, S., Ma, H., Zhang, L., and Sun, H. (2020). *A review for cervical histopathology image analysis using machine vision approaches*, volume 53.

Li, C., Chen, H., Xue, D., Hu, Z., Zhang, L., He, L., Xu, N., Qi, S., Ma, H., and Sun, H. (2019a). Weakly Supervised Cervical Histopathological Image Classification Using Multilayer Hidden Conditional Random Fields. In *Advances in Intelligent Systems and Computing*, volume 1011, pages 209–221. Springer Verlag.

Li, C., Chen, H., Zhang, L., Xu, N., Xue, D., Hu, Z., Ma, H., and Sun, H. (2019b). Cervical Histopathology Image Classification Using Multilayer Hidden Conditional Random Fields and Weakly Supervised Learning. *IEEE Access*, 7:90378–90397.

Litjens, G., Bandi, P., Bejnordi, B. E., Geessink, O., Balkenhol, M., Bult, P., Halilovic, A., Hermsen, M., van de Loo, R., Vogels, R., Manson, Q. F., Stathonikos, N., Baidoshvili, A., van Diest, P., Wauters, C., van Dijk, M., and van der Laak, J. (2018). 1399 H&E-stained sentinel lymph node sections of breast cancer patients: The CAMELYON dataset. *GigaScience*, 7(6):1–8.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 2, pages 1150–1157. IEEE.

Ninos, K., Kostopoulos, S., Kalatzis, I., Sidiropoulos, K., Ravazoula, P., Sakellaropoulos, G., Panayiotakis, G., Economou, G., and Cavouras, D. (2016). Microscopy image analysis of p63 immunohistochemically stained laryngeal cancer lesions for predicting patient 5-year survival. *European Archives of Oto-Rhino-Laryngology*, 273(1):159–168.

Ouddai, G., Hamdi, I., and Ben Ghezala, H. (2023). Color-SIFT Features for Histopathological Image Analysis. In Lee, R., editor, *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing. SNPD 2022. Studies in Computational Intelligence. (in press)*. Springer cham.

Öztürk, Å and Akdemir, B. (2018). Application of Feature Extraction and Classification Methods for Histopathological Image using GLCM, LBP, LBGLCM, GLRLM and SFTA. In *Procedia Computer Science*, volume 132, pages 40–46. Elsevier B.V.

Popovici, V., Budinská, E., Čápková, L., Schwarz, D., Dušek, L., Feit, J., and Jaggi, R. (2016). Joint analysis of histopathology image features and gene expression in breast cancer. *BMC Bioinformatics*, 17(1):1–9.

Rosten, E., Porter, R., and Drummond, T. (2010). Faster and better: A machine learning approach to corner detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(1):105–119.

Sanchez-Morillo, D., González, J., García-Rojo, M., and Ortega, J. (2018). Classification of Breast Cancer Histopathological Images Using KAZE Features. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10814 LNBI:276–286.

Shafiei, S., Babaie, M., Kalra, S., and Tizhoosh, H. R. (2021). Colored Kimia Path24 Dataset: Configurations and Benchmarks with Deep Embeddings.

Simon, O., Yacoub, R., Jain, S., Tomaszewski, J. E., and Sarder, P. (2018). Multi-radial LBP Features as a Tool for Rapid Glomerular Detection and Assessment in Whole Slide Histopathology Images. *Scientific Reports*, 8(1):1–11.

Spanhol, F. A., Oliveira, L. S., Petitjean, C., and Heutte, L. (2016). A Dataset for Breast Cancer Histopathological Image Classification. *IEEE Transactions on Biomedical Engineering*, 63(7):1455–1462.

Tola, E., Lepetit, V., and Fua, P. (2010). DAISY: An efficient dense descriptor applied to wide-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(5):815–830.

Wang, C. W., Chang, C. C., Khalil, M. A., Lin, Y. J., Liou, Y. A., Hsu, P. C., Lee, Y. C., Wang, C. H., and Chao, T. K. (2022). Histopathological whole slide image dataset for classification of treatment effectiveness to ovarian cancer. *Scientific Data*, 9(1):1–5.