Deep Ordinal Focus Assessment for Whole Slide Images

Tomé Albuquerque
INESC TEC
FEUP
tome.m.albuquerque@inesctec.pt

Ana Moreira
FEUP

Jaime S. Cardoso
INESC TEC
FEUP
jaime.cardoso@inesctec.pt

Abstract

Medical image quality assessment plays an important role not only in the design and manufacturing processes of image acquisition but also in the optimization of decision support systems. This work introduces a new deep ordinal learning approach for focus assessment in whole slide images. From the blurred image to the focused image there is an ordinal progression that contains relevant knowledge for more robust learning of the models. With this new method, it is possible to infer quality without losing ordinal information about focus since instead of using the nominal cross-entropy loss for training, ordinal losses were used. Our proposed model is contrasted against other state-of-the-art methods present in the literature. A first conclusion is a benefit of using data-driven methods instead of knowledge-based methods. Additionally, the proposed model is found to be the top-performer in several metrics. The best performing model scores an accuracy of 94.4% for a 12 classes classification problem in the FocusPath database.

1. Introduction

In the last few decades, there was a massive growth in the amount of digital color image content in the medical field due to the spread of advanced multimedia devices and digital services capable of doing acquisition, transmission, and storage of digital data [8]. Digital Pathology images (Whole slide images (WSI)) are about 10X bigger than Radiology images, being over $>$1 GB in size in most cases, this type of image requires better storage management through their useful life cycle in clinical workflow [7, 30]. Thus, image quality assessment (IQA) methods are crucial, working as a filter in the first stage of acquisition, they can improve storage management. In a second step, they can be used as an optimizer of the decision support systems.

In the medical field image quality assessment methodologies are focused on two distinct processes: on the one hand, the low-level notion of quality which includes the measurement of distortions at a signal level such as blur, noise, compression errors, and other types of distortions [10]; on the other hand, the semantic complex concepts such as the presence/absence of artifacts (e.g. tissue folds or bubbles, the presence of coloration errors, among others) [3]. Thus, the medical image quality assessment is, in most cases, application-specific which requires vast domain knowledge in the respective medical area.

Focus quality assessment (FQA) is fundamental in the normal WSIs acquisition workflow. Hence, it is essential to develop and improve FQA methods capable of improving the quality of acquired WSIs and reduce the acquisition time.

In the next Section 2, other work for focus quality assessment in WSIs is reviewed. A brief summary of the problem and the contributions of this work is done in subsection 3 The Section 4 describes the Data Pre-Processing steps, and it also presents a brief introduction to Convolutional Neural Networks, the used Network Architectures, and finally, the losses used during the training. In Section 5 the experimental details are described followed by the results and discussion in Section 6. The study finished with a Conclusions Section 7.

2. Related Work

During the acquisition of WSI's by the scanning platforms, focus errors often occur. After the acquisition, manual inspection of the slides is required to infer if it has good quality to proceed with the analysis and diagnostic. The manual inspection of the slides is a time-consuming process and in most cases subjective to individual scores which leads to inter/intra-variability issues between the experts [29]. This way, to automatize this process and improve clinical workflow several focus quality assessment approaches have been developed. They can be divided into Knowledge-based focus quality assessment methods and data-driven focus quality assessment methods.

Knowledge-based FQAs - These methods are based on domain knowledge, with a large presence of this type of algorithms in the literature. Knowledge-based FQAs can be divided into 7 different categories: Derivative-Based Al-

gorithms [32, 39], Statistics-Based Algorithms [26, 2, 13], Histogram-Based Algorithms [11], Intuitive Algorithms (based on thresholds) [28], microscopic optic models [16], signal processing models [25, 14] and also human visual system models [17, 18]. Most of the methods previously mentioned methods require low computational power and are more interpretable; however, compared with the most recent data-driven methods, their performance is relatively low in terms of precision and computation time.

MLV [2] and FQPATH [16] are two state-of-the-art knowledge-based models in the literature. The MLV method quantifies the image quality considering the maximum variation between each pixel and its eight neighbours. The final score is given by the standard deviation of the values found for each pixel, after applying a weighting function, which attributes more relevance to higher values, increasing the difference of scores between focus and defocus images. The FQPath method adopts a kernel with properties similar to the human visual system as a sum of even derivative filters. This kernel is then applied to the image, outputting a feature map that is vectorized. Then only the values coding the most focus-relevant features are preserved, ignoring features related with noise artifacts. Lastly the m-th central moment of the remaining features is calculated and the final score is given by the negative logarithm of this value. The optimal value of m can be found through a grid search.

Data-Driven FQAs - In the last few years, there was an advance in data-driven approaches for FQA in WSIs. These type of methods require a large number of data for the training, however, presents in most of the cases a boost in the performance when compared to Knowledge-based FQAs. Inside these types of models, literature presents two different categories: network adapted with minor adjustments from pre-designed and well-known architectures [23, 5], and on the other hand, architectures tailored from scratch with application-specific features [37, 31, 33, 38]. There are some open-source software applications available in the literature that try to apply some approaches for FQA in WSIs derived from the methods mentioned above: HistoQC [22] and CellPro [27]. Data-Driven approaches present very good performances in WSIs focus assessment, however, the high computational costs to train these models represent its main drawbacks.

3. Brief Summary

Most of the current approaches to ordinal inference for neural networks are found to not adequately take advantage of the ordinal problem. In the case of FQA for WSIs, all the data-driven approaches present in the literature discard the ordinal information between the different focus classes.

In the present work, a novel model to infer focus quality in WSIs is proposed. This new model is developed on the FocusPath dataset and uses a plethora of deep learning architectures, with several ordinal losses. Our proposal is also compared with the current state-of-the-art approaches and surpasses them in several metrics.

4. Methods

4.1. Data pre-processing

Firstly, the dataset was divided into train, validation, and test subsets (60-20-20%), maintaining the ratio among different classes. To feed the network, it was necessary to resize the patch images from 1024×1024px to 224×224px. This process was done online during the training. It was also done normalization to speed up the training, which consists of a rescale of pixel values from the range of 0-255 to 0-1. To overcome the small amount of data in our dataset, data augmentation was used. Therefore, during the training of the models, a series of random transformations were done in each training epoch for every image. The transformations applied included 10% of width and height shift, 10% of zoom, horizontal and vertical flips, and image rotation. Figure 1 shows three different examples of random transformations that were done online during the training process of the models.

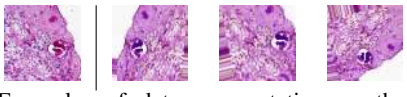

Figure 1. Examples of data augmentation on the FocusPath database. From left to right: original image and three examples of random transformations

4.2. Convolutional Neural Networks

A convolutional neural network (CNN) is a class of deep learning algorithms that consecutively apply convolutions of filters to the image. These filters are learned and consist of quadrilateral patches that are convolved across the whole input image – unlike previous fully-connected networks, only local inputs are connected at each layer. Usually, each convolution is intertwined with downsampling operations, such as max-pooling, that progressively reduce the size of the original input image.

The final layers are fully-connected and then the final output is processed by a soft-max for the multi-class problems. Dropout was used to reduce overfitting by constraining these fully-connected layers [35]. This type of model has the intrinsic ability to learn useful features directly from the input images that are important for the task at hand.

4.3. Network Architectures

In this work a model for classification of focus quality in WSIs was trained and tested with seven different convolutional network architectures: AlexNet [24],

GoogLeNet [36], ResNet18 [15], MobileNet_V2 [20], ShuffleNet_V2_X1_0 [40], SqueezeNet1_0 [21], and VGG-16 [34]. The main goal of testing these different architectures is to compare the performance of a non-ordinal loss (CE) vs ordinal losses in a wide range of architectures. These seven different architectures were chosen as they are well known and often used in the literature. They came pre-trained with PyTorch on ImageNet[1].

The last block of each architecture was replaced by the following layers: dropout with p=20%, 512-unit dense layer with ReLU, dropout with p=20%, a 256-wide dense layer with ReLU, followed by 12 output neurons.

4.4. Losses

In this work six different losses will be evaluated, Cross-Entropy (CE) for the baseline model and five different ordinal losses: Ordinal Encoding (OE), Binomial Unimodal (BU), CO, CO2, and also Ordinal Entropy Loss Function (HO2).

Cross-Entropy (CE): One of the most used losses in deep learning problems, where a model perform multi-class classification by minimizing cross-entropy (equation 1), averaged over the training set:

$$\text{CE}(\mathbf{y}_n, \hat{\mathbf{y}}_n) = -\sum_{k=1}^{K} y_{nk} \log(\hat{y}_{nk}) \tag{1}$$

Cross-entropy is a suitable approach for nominal data. Nonetheless, for ordinal data, where the order between classes gives relevant insights, this information must be explored by the loss to further regularize learning.

Ordinal Encoding (OE): In ordinal encoding, which is an agnostic way to introduce ordinality, classes are encoded using a cumulative distribution – the indicator function used is $\mathbb{1}(k < k^\star)$ so that y_{nm} is represented by 1 if $k < k_n^\star$ and 0 otherwise. Each output represents the incremental neighbor probability, and the inverse operation (during inference) is performed by summing up these outputs, $p_{nk} = \sum_{m=1}^{K-1} y_{nm}$ [12, 6].

Binomial Unimodal (BU): Constraining discrete ordinal probability distributions to be unimodal using binomial probability distributions is another technique to promote ordinality in classification problems. Thus, it is necessary to constrain the output of the network directly under a regression setting. This method predicts a single output (as the final layer) representing the probability along the classes, with $y_n = 0$ representing $k_n^\star = 1$ and $y_n = 1$ representing $k_n^\star = K$ [9, 4].

CO and CO2 Ordinal losses - In this loss, a regularization term that penalizes the deviations from the unimodal setting is added to CE [1].

[1] https://pytorch.org/docs/stable/torchvision/models.html

Defining $\mathbb{1}(x)$ as the indicator function of x and $\text{ReLU}(x) = x\mathbb{1}(x > 0) = \max(0, x)$, a possible fix for an order-aware loss could be

$$\begin{aligned}\text{CO}(\mathbf{y}_n, \hat{\mathbf{y}}_n) &= \text{CE}(\mathbf{y}_n, \hat{\mathbf{y}}_n) \\ &+ \lambda \sum_{k=1}^{K-1} \mathbb{1}(k \geq k_n^\star)\, \text{ReLU}(\hat{y}_{n(k+1)} - \hat{y}_{n(k)}) \\ &+ \lambda \sum_{k=1}^{K-1} \mathbb{1}(k \leq k_n^\star)\, \text{ReLU}(\hat{y}_{n(k)} - \hat{y}_{n(k+1)}),\end{aligned} \tag{2}$$

where $\lambda \geq 0$ controls the relative influence of the extra terms favoring unimodal distributions.

The additional terms, although promoting uni-modality, still allow flat distributions. A generalization of the previous idea is to add a margin of $\delta > 0$ to the ReLU, imposing that the difference between consecutive probabilities is at least δ [1]. This leads us to a second CE loss, **CO2**, suitable for ordinal classes:

$$\begin{aligned}\text{CO2}(\mathbf{y}_n, \hat{\mathbf{y}}_n) &= \text{CE}(\mathbf{y}_n, \hat{\mathbf{y}}_n) \\ &+ \lambda \sum_{k=1}^{K-1} \mathbb{1}(k \geq k_n^\star)\, \text{ReLU}(\delta + \hat{y}_{n(k+1)} - \hat{y}_{n(k)}) \\ &+ \lambda \sum_{k=1}^{K-1} \mathbb{1}(k \leq k_n^\star)\, \text{ReLU}(\delta + \hat{y}_{n(k)} - \hat{y}_{n(k+1)}).\end{aligned} \tag{3}$$

A value of $\delta = 0.05$ has been empirically found to provide a sensible margin.

Ordinal Entropy Loss Function (HO2) This loss arises to mitigate some negative aspects of the CE in the CO and CO2 losses since the CE term by itself is only trying to maximize the probability estimated in the true output class (while ignoring the remaining probabilities). The ordinal terms in CO and CO2 are promoting unimodality but not penalizing (almost) flat distributions. A softer assumption is that the distribution should have a low entropy [1].

This leads us to the ordinal entropy loss, HO2, for ordinal data as

$$\begin{aligned}\text{HO2}(\mathbf{y}_n, \hat{\mathbf{y}}_n) &= \text{H}(\hat{\mathbf{y}}_n) \\ &+ \lambda \sum_{k=1}^{K-1} \mathbb{1}(k \geq k_n^\star)\, \text{ReLU}(\delta + \hat{y}_{n(k+1)} - \hat{y}_{n(k)}) \\ &+ \lambda \sum_{k=1}^{K-1} \mathbb{1}(k \leq k_n^\star)\, \text{ReLU}(\delta + \hat{y}_{n(k)} - \hat{y}_{n(k+1)}),\end{aligned} \tag{4}$$

where $\text{H}(\mathbf{p})$ denotes the entropy of the distribution $\mathbf{p}$.

5. Experimental Details

5.1. Dataset

In this work, it was used a public histopathological database named FocusPath dataset [2], which contains image quality annotations [19]. The FocusPath dataset contains 8640 patches of 1024 x 1024 images. These images were extracted from nine different stained slides from diverse human organs. The original Whole Slide Images were scanned by Huron TissueScope LE1.2. It uses a 40X optics lens at 0.25 μm/pixel resolution. There are 14 absolute z-level scores corresponding to the ground-truth class for the focus level. In Table 1 it is possible to see several examples from each class present in the FocusPath dataset. Due to the low number of examples of images in the more defocused classes (12 and 13), the label was changed to belong to class 11. This way, the dataset was grouped into 12 different focus classes.

5.2. Training

All the losses were tested on the CNN model represented on Figure 2.

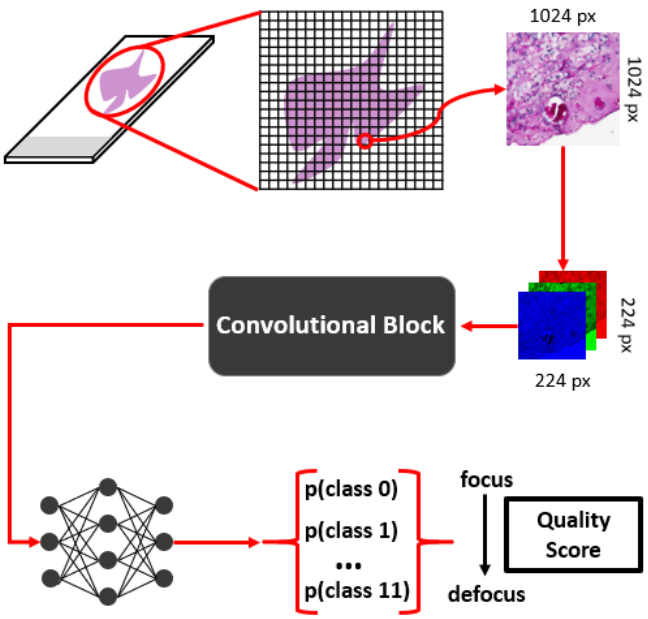

Figure 2. Schematic representation of the model architecture.

The model had as input the RGB WSIs patches from the FocusPath database. The model output consists of the multiclass classification (12 classes) of focus quality (0-focus patch to 11-defocus patch)

During the training initialization, the weights of the architectures previously mentioned were initialized based on ImageNet pre-training. The optimizer used was ADAM, starting with a learning rate of 10^{-4}. The learning rate is reduced by 10% whenever the loss is stagnant for 10 epochs using a specific scheduler. The training process is completed after 200 epochs. The training process was conducted on a single Nvidia GTX 1080ti (11GB) GPU.

In the case of the Ordinal Entropy Loss Function, the hyperparameter λ is tuned by doing nested k-fold cross-validating using the training set (with k=5) to create an unbiased validation set.

5.3. Evaluation Metrics

To evaluate the performance of the different models, three different metrics were adopted: accuracy, mean absolute error, and also Kendall's τ rank correlation coefficient.

One of the most popular metrics used in classification problems is accuracy (Acc). For N observations, taking k_i and $\hat{k}_i$ to be the label and prediction of the n-th observation, respectively, then Acc $= \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}(\hat{k}_n^\star = k_n^\star)$, where $\mathbb{1}$ is the indicator function.

However, this metric treats all class errors as equal, whether the error is between adjacent classes or between classes in the extreme. If we have K classes represented by a set $\boldsymbol{\mathcal{C}} = \{\mathcal{C}^{(1)}, \mathcal{C}^{(2)}, \ldots, \mathcal{C}^{(K)}\}$, then accuracy will treat an error between $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(2)}$ with the same magnitude as an error between $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(K)}$ which is clearly worse. For that reason, a popular metric for ordinal classification is the Mean Absolute Error (MAE), MAE $= \frac{1}{N}\sum_i |k_i^\star - \hat{k}_i^\star|$. This metric is not perfect since it treats an ordinal variable as a cardinal variable. An error between classes $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(3)}$ will be treated as two times worse than an error between classes $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(2)}$. Naturally, the assumption of cardinality is not always warranted.

Kendall's τ rank correlation coefficient (Kendal Tau) is a non-parametric evaluation of relationships between columns of ranked data, so it is a measure of ordinal association between data. The τ correlation coefficient returns a value that ranges from -1 to 1, with 0 being no correlation and 1 perfect correlation.

6. Results and Discussion

The performance of the seven different architectures are presented in Tables 2–3, and 4, for 12-class classification problem, with the six different learning losses – conventional Cross-Entropy (CE), Binomial Unimodal (BU) [9], Ordinal Encoding (OE) [6], CO, CO2 and Ordinal Entropy Loss Function (HO2) [1]. The best models are shown in bold.

Table 2 shows the accuracy results for 12-class focus classification problem. Regarding this metric, it is possible to infer the difference in the results between nominal and ordinal losses. OE loss performance was better for 4 different architectures against 2 for nominal CE. This can be explained by the lower role of ordinality in the CE loss, as

[2] https://zenodo.org/record/3926181#.YPFgluhKjIU

Table 1. Image examples of the 14 Different patch focus Classes in the FocusPath dataset.

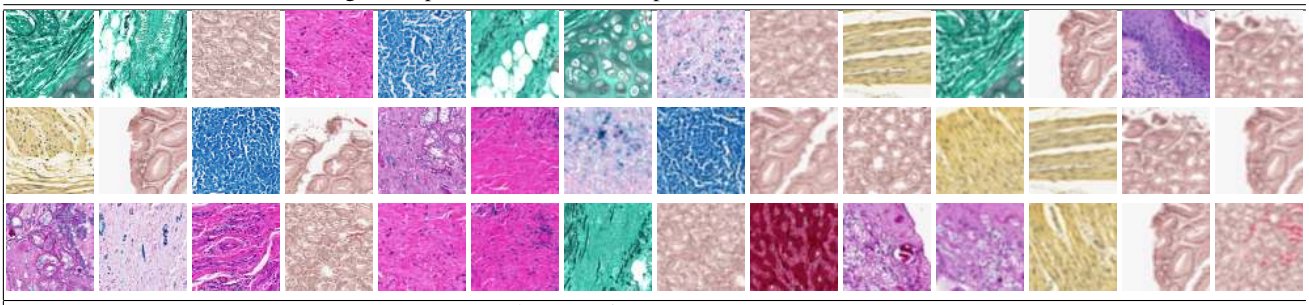

Table 2. Results in terms of **Accuracy (%)** for **12 class** problem (higher is better).

	CE	BU	OE	CO	CO2	HO2
alexnet	**72.4**	69.7	66.8	64.9	62.7	62.9
googlenet	92.2	87.2	91.1	**92.4**	87.0	78.1
resnet18	**88.7**	86.8	87.2	83.3	82.7	86.6
mobilenet_v2	92.2	82.9	**94.4**	89.3	86.0	86.0
shufflenet_v2_x1_0	84.7	81.9	**86.0**	84.3	76.1	80.8
squeezenet1_0	73.4	71.5	**74.2**	31.1	63.9	62.9
vgg16	87.2	84.9	**89.1**	62.1	84.3	80.6
Avg	84.4	80.7	84.1	72.5	77.5	76.8
Winners	2	0	**4**	1	0	0

Table 3. Results in terms of **Mean Absolute Error (MAE)** for **12 class** problem (lower is better).

	CE	BU	OE	CO	CO2	HO2
alexnet	**0.29**	0.32	0.34	0.37	0.42	0.40
googlenet	0.08	0.13	0.09	**0.08**	0.13	0.22
resnet18	**0.12**	0.13	0.13	0.19	0.19	0.14
mobilenet_v2	0.08	0.18	**0.06**	0.12	0.15	0.16
shufflenet_v2_x1_0	0.16	0.19	**0.16**	0.16	0.25	0.20
squeezenet1_0	0.30	0.29	**0.27**	1.04	0.43	0.42
vgg16	0.14	0.15	**0.11**	0.43	0.17	0.23
Avg	0.17	0.20	0.17	0.34	0.25	0.25
Winners	2	0	**4**	1	0	0

Table 4. Results in terms of **Kendall's τ** for **12 class** problem (higher is better).

	CE	BU	OE	CO	CO2	HO2
alexnet	**93.8**	93.5	93.3	92.2	91.5	91.7
googlenet	**98.4**	97.2	98.0	98.3	97.3	95.7
resnet18	97.4	97.1	**97.4**	95.8	96.0	96.8
mobilenet_v2	98.3	96.6	**98.8**	97.5	96.9	96.2
shufflenet_v2_x1_0	96.3	95.8	96.3	**96.3**	95.0	96.0
squeezenet1_0	92.8	**94.2**	94.1	75.0	90.7	91.3
vgg16	97.2	97.1	**97.9**	89.9	96.2	94.8
Avg	96.3	95.9	96.5	92.1	94.8	94.7
Winners	2	1	**3**	1	0	0

also confirmed by the MAE results. This means that when misclassification occurs, ordinal losses tend to classify the focus quality in the WSI patches as being closer to the real class.

Table 3 presents the results for MAE, which confirm the influence of ordinal losses in promoting ordinality when comparing to nominal loss (CE). OE loss achieved the best performance across the different architectures. Across the different architectures, ordinal losses won nominal CE in 5 against 2.

In Figure 3 is presented the confusion matrix of our best model (MobileNet_v2 trained with Ordinal Encoding). It is possible to infer that the misclassified images were very close to the main diagonal of the confusion matrix, which reveals that the ordinal loss helps to reduce the distance between the true class vs predicted class when the model classifies wrongly as mentioned before.

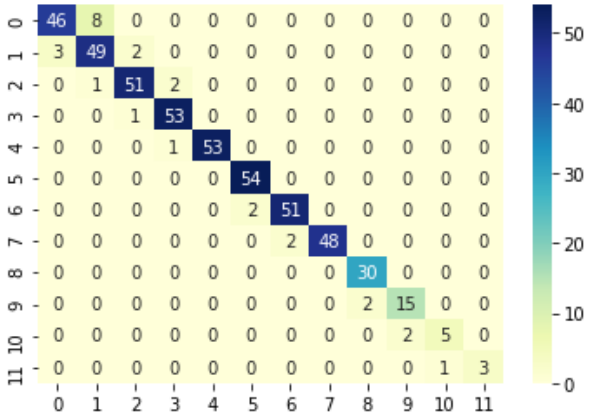

Figure 3. Confusion Matrix of our best model - **MobileNet_v2** - using Ordinal Encoding as loss.

Going a step further and comparing the results with other works featuring focus quality assessment in the FocusPath database, it is clear that the evaluation metrics are above the state-of-the-art results, especially on the Spearman's rank correlation coefficient (SRCC) and Pearson correla-

tion coefficient (PLCC) metrics. Wang et.al [37] present a data-driven work with better results in all the metrics, in contrast to knowledge-based methods (MLV [2] and FQ-PATH [16]). However, in his approach, the model is fed with random crops (235 x 235 px) of the WSIs patches. Thus, it loses some relevant spatial information about the focus (e.g. when exists in the same patch focus and defocus zones).

This comparison of metrics between our proposed method and the literature works is presented in Table 5.

Table 5. Comparison of our proposal with literature models

	SRCC	PLCC	Time (sec)
MLV [2]	0.8623	0.8528	0.482
FQPATH [16]	0.8395	0.8295	0.269
FOCUSLITENN [37]	0.8931	0.8857	**0.019**
Proposal	**0.9969**	**0.9970**	0.047

7. Conclusions and Future Work

Comparing the different deep learning approaches on WSIs patches, the models trained with ordinal losses achieved better results when comparing with the nominal cross-entropy loss. Thus, a new model has been proposed for multi-class FQA in WSIs based on convolutional neural networks and respecting the ordinal progression among the different focus levels. The model was trained in a dataset with several color spectrums and tissue structures, these factors contributed to a robust training of the model. This new model demonstrated to be competitive with state-of-the-art results and surpass them in some metrics.

Furthermore, another important outcome is that those data-driven methods obtained significantly better results than knowledge-based methods in terms of precision and performance.

For future works, since this model is a prototype of an FQA system, the architecture should suffer some changes and multi-processing must be used to reduce the processing time per patch in a real-time acquisition system for WSIs.

Acknowledgments

The project TAMI-Transparent Artificial Medical Intelligence (NORTE-01-0247-FEDER-045905) partially funding this work is co-financed by ERDF - European Regional Fund through the North Portugal Regional Operational Program-NORTE 2020 and by the Portuguese Foundation for Science and Technology-FCT under the CMU-Portugal International Partnership. Tomé Albuquerque was supported by Ph.D. grant 2021.05102.BD, also provided by FCT.

References

[1] Tomé Albuquerque, Ricardo Cruz, and Jaime S. Cardoso. Ordinal losses for classification of cervical cancer risk. *PeerJ Computer Science*, 7:e457, Apr. 2021.

[2] Khosro Bahrami and Alex C. Kot. A fast approach for no-reference image sharpness assessment based on maximum local variation. *IEEE Signal Processing Letters*, 21(6):751–755, 2014.

[3] Pinky A. Bautista and Yukako Yagi. Detection of tissue folds in whole slide images. In *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009*, 2009.

[4] Christopher Beckham and Christopher Pal. Unimodal probability distributions for deep ordinal classification. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 411–419, International Convention Centre, Sydney, Australia, 06–11 Aug 2017. PMLR.

[5] Gabriele Campanella, Arjun R Rajanna, Lorraine Corsale, Peter J Schüffler, Yukako Yagi, and Thomas J Fuchs. Towards machine learned quality control: A benchmark for sharpness quantification in digital pathology. *Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society*, 65:142–151, apr 2018.

[6] Jianlin Cheng, Zheng Wang, and Gianluca Pollastri. A neural network approach to ordinal regression. In *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, pages 1279–1284. IEEE, 2008.

[7] Elizabeth Chlipala, Jesus Elin, Ole Eichhorn, Richard E Long, Bikash Sabata, and Matthew Smith. Archival and Retrieval in Digital Pathology Systems. *Digital Pathology Association*, 2011.

[8] Cisco. Cisco VNI forecast and methodology, 2014-2019. *Cisco Visual Networking Index*, 2015.

[9] Joaquim Costa and Jaime Cardoso. Classification of ordinal data using neural networks. In *European Conference on Machine Learning*, pages 690–697, 2005.

[10] Yong Ding. *Visual Quality Assessment for Natural and Medical Image*. Springer, 03 2018.

[11] Lawrence Firestone, Kitty Cook, Kevin Culp, Neil Talsania, and Kendall Preston. Comparison of autofocus methods for automated microscopy. *Cytometry*, 1991.

[12] Eibe Frank and Mark Hall. A simple approach to ordinal classification. In *European Conference on Machine Learning*, pages 145–156, 2001.

[13] Frans C.A. Groen, Ian T. Young, and Guido Ligthart. A comparison of different focus functions for use in autofocus algorithms. *Cytometry*, 1985.

[14] Rania Hassen, Zhou Wang, and Magdy M. A. Salama. Image sharpness assessment based on local phase coherence. *IEEE Transactions on Image Processing*, 22(7):2798–2810, 2013.

[15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceed-*

ings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.

[16] M. S. Hosseini, Jasper A. Z. Brawley-Hayes, Y. Zhang, Lyndon Chan, K. Plataniotis, and S. Damaskinos. Focus quality assessment of high-throughput whole slide imaging in digital pathology. *IEEE Transactions on Medical Imaging*, 39:62–74, 2020.

[17] Mahdi S. Hosseini and Konstantinos N. Plataniotis. Image sharpness metric based on maxpol convolution kernels. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 296–300, 2018.

[18] Mahdi S. Hosseini, Yueyang Zhang, and Konstantinos N. Plataniotis. Encoding visual sensitivity by maxpol convolution filters for image sharpness assessment. *IEEE Transactions on Image Processing*, 28(9):4510–4525, 2019.

[19] M. S. Hosseini, Y. Zhang, and K. N. Plataniotis. Encoding visual sensitivity by maxpol convolution filters for image sharpness assessment. *IEEE Transactions on Image Processing*, 28(9):4510–4525, Sep. 2019.

[20] Andrew G Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017.

[21] Forrest N Iandola, Song Han, Matthew W Moskewicz, Khalid Ashraf, William J Dally, and Kurt Keutzer. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and ¡0.5 MB model size. *arXiv preprint arXiv:1602.07360*, 2016.

[22] Andrew Janowczyk, Ren Zuo, Hannah Gilmore, Michael Feldman, and Anant Madabhushi. HistoQC: An Open-Source Quality Control Tool for Digital Pathology Slides. *JCO clinical cancer informatics*, 3:1–7, apr 2019.

[23] Timo Kohlberger, Yun Liu, Melissa Moran, Po-HsuanCameron Chen, Trissia Brown, JasonD Hipp, CraigH Mermel, and MartinC Stumpe. Whole-slide image focus quality: Automatic assessment and impact on ai cancer detection. *Journal of Pathology Informatics*, 10(1):39, 2019.

[24] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. ImageNet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[25] Arthur Leclaire and Lionel Moisan. No-reference image quality assessment and blind deblurring with sharpness metrics exploiting fourier phase information. *Journal of Mathematical Imaging and Vision*, 52:145–172, 05 2015.

[26] Leida Li, Dong Wu, Jinjian Wu, Haoliang Li, Weisi Lin, and Alex C. Kot. Image sharpness assessment by sparse representation. *IEEE Transactions on Multimedia*, 18(6):1085–1097, 2016.

[27] Claire McQuin, Allen Goodman, Vasiliy Chernyshev, Lee Kamentsky, Beth A. Cimini, Kyle W. Karhohs, Minh Doan, Liya Ding, Susanne M. Rafelski, Derek Thirstrup, Winfried Wiegraebe, Shantanu Singh, Tim Becker, Juan C. Caicedo, and Anne E. Carpenter. Cellprofiler 3.0: Next-generation image processing for biology. *PLOS Biology*, 16(7):1–17, 07 2018.

[28] Mortimer L. Mendelsohn and Brian H. Mayall. Computer-oriented analysis of human chromosomes—iii. focus. *Computers in Biology and Medicine*, 2(2):137–150, 1972. Chromosome Analysis.

[29] Xavier Moles Lopez, Etienne D'Andrea, Paul Barbot, Anne-Sophie Bridoux, Sandrine Rorive, Isabelle Salmon, Olivier Debeir, and Christine Decaestecker. An automated blur detection method for histological whole slide imaging. *PloS one*, 8:e82710, 12 2013.

[30] Sara Oliveira, Pedro C. Neto, Joao Fraga, Diana Montezuma, Ana Monteiro, Joao Monteiro, Liliana Ribeiro, Sofia Goncalves, Isabel M. Pinto, and Jaime S. Cardoso. Cad systems for colorectal cancer from wsi are still not ready for clinical acceptance. *Scientific Reports (Nature)*, 2021.

[31] Henry Pinkard, Zachary Phillips, Arman Babakhani, Daniel A. Fletcher, and Laura Waller. Deep learning for single-shot autofocus microscopy. *Optica*, 6(6):794–797, Jun 2019.

[32] A Santos, Carlos Ortiz-de Solorzano, Juan Jose Vaquero, J Peña, Norberto Malpica, and Francisco Del Pozo Guerrero. Evaluation of autofocus functions in molecular cytogenetic analysis. *Journal of microscopy*, 188:264–72, 01 1998.

[33] Caglar Senaras, M Khalid Khan Niazi, Gerard Lozanski, and Metin N Gurcan. DeepFocus: Detection of out-of-focus regions in whole slide digital images using deep learning. *PloS one*, 13(10):e0205387–e0205387, oct 2018.

[34] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[35] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15:1929–1958, 06 2014.

[36] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

[37] Zhongling Wang, Mahdi S. Hosseini, Adyn Miles, Konstantinos N. Plataniotis, and Zhou Wang. Focuslitenn: High efficiency focus quality assessment for digital pathology, 2020.

[38] Samuel J Yang, Marc Berndl, D Michael Ando, Mariya Barch, Arunachalam Narayanaswamy, Eric Christiansen, Stephan Hoyer, Chris Roat, Jane Hung, Curtis T Rueden, Asim Shankar, Steven Finkbeiner, and Philip Nelson. Assessing microscope image focus quality with deep learning. *BMC Bioinformatics*, 19(1):77, 2018.

[39] TTE Yeo, SH Ong, Jayasooriah, and R Sinniah. Autofocusing for tissue microscopy. *Image and Vision Computing*, 11(10):629–639, 1993.

[40] Xiangyu Zhang, Xinyu Zhou, Mengxiao Lin, and Jian Sun. ShuffleNet: An extremely efficient convolutional neural network for mobile devices. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 6848–6856, 2018.