Fine-Grained Deep Learning for Gleason Grading in Prostate Histopathology

Kaixin Chen

College of Computer Science, Sichuan University, Chengdu Sichuan 610207, China
ckx1216@outlook.com

Abstract. One of the most frequent cancers in male patients is prostate cancer, diagnosis and prognosis of which include histopathological image analysis, especially Gleason grading. It is a delicate activity that involves identifying subtle architectural variations in tissue patterns which are however disadvantaged by inter-observer imprecision, subjectiveness, and labor intensity. Deep learning has become a powerful solution, offering an automated, reproducible grading architecture modelling intra-class differences and visual similarities. This survey discusses recent state-of-the-art deep learning methods in the field of fine-grained prostate cancer histopathology, which are whole-slide image analysis and Gleason pattern recognition. It studies convolutional neural networks, attention mechanisms and vision transformers with regard to their capability of localizing diagnostic regions and multi-scale context integration. The weakly supervised learning and transfer learning, as well as interpretation techniques designed to minimize the annotation need and promote clinical trust, are also discussed in the paper. Such benchmark datasets as The Cancer Genome Atlas (TCGA) and measures as quadratic Cohen's kappa and area under the curve (AUC) are examined. Continuous problems are the staining discrepancy, unclear labels, and computational needs. Prospective clinical validation remains crucial for future translation.

Keywords: Fine-grained Image Classification, Deep Learning, Prostate Cancer, Histopathological Images, Gleason Grading.

1 Introduction

Fine-grained learning has emerged as a pivotal paradigm in the automated analysis of histopathological images, particularly for prostate cancer (PCa), which remains one of the most prevalent malignancies worldwide and a leading cause of cancer-related mortality [1]. Gleason grading system, which is the foundation of PCa prognosis, is based on the architectural distinction between malignant epithelial patterns seen in stained tissue section, but it is a subjective, labor intensive, and restricted method of assessment by inherent inter-interpreter variance among pathologists [2]. To mitigate these limitations, deep learning (DL)-driven computer-aided diagnosis (CAD) systems have been increasingly explored to enhance both diagnostic accuracy and reproducibility [2]. Recent developments show that high-precision predictions like

K. Subramanian (ed.), *Proceedings of the International Workshop on Advances in Deep Learning for Image Analysis and Computer Vision (IWADIC 2025)*, Advances in Computer Science Research 128,
https://doi.org/10.2991/978-94-6239-648-7_64

0.98 ± 0.02 Gleason score predictions through the use of shearlet transforms and expert-informed saliency maps using a combination of Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and fine-tuned VGGNet have been demonstrated through recent progress [1]. Simultaneously, other imaging techniques like multiphoton microscopy (MPM) which allows visualization of unstained tissues with label-free and subcellular-resolution have also been adopted to classifiers using AlexNet with an area under the curve (AUC) of 0.943, thus denoting high levels of discriminative features in automated grading [2]. In addition, to overcome the typical difficulties of fine-grained classification, including small inter-class variability and large intra-class variation, particularly at different magnifications, new frameworks by adding multi-task learning and embedding the features without being influenced by magnification, have proven to generalize better over breast cancer histopathology data, which could be transferred to analysing PCa [1]. It is important to note, though, that, in addition to pixel-level or patch-based methods, there are topological cellular network features, such as cycle structures in spatial cell graphs, which have been demonstrated to have better performance than conventional texture- and fractal-based measures when differentiating between Gleason grade 3 and grade 4 tumors by using Support Vector Machine (SVM) classifiers [1]. The above developments are consistent with the larger digital pathology trends, with applications in this field being largely diagnostic (including detecting and grading cancer), with certain applications reaching AUCs of nearly 0.99 [2]. Notably, the emerging deep learning systems (DLS) do not only recapitulate but also optimise the grading process providing probabilistic, fine-grained pattern prediction, which demonstrates the histological ambiguity in transitional morphological boundaries and, therefore, allows the risks to be stratified more than ever before [3][4][5]. In that way, the path of integrating fine-grained categorization into clinical workflows has potentials of decreasing diagnostic variability and enhancing prognostic accuracy, especially now that the large-scale, fine-grained benchmark further develops [6].

This survey conducts a systematic review on the subject of fine-grained learning approaches, which are specifically designed to be used in the domain of prostate cancer histopathological image analysis, and more precisely, the recent developments in the field of deep learning. It sets the scene by explaining the clinical significance of Gleason grading which is a standardized diagnostic tool used in prostate cancer to aid in its prognosis [6]. The article emphasizes that deep learning models composed of hierarchical transformer-based neural networks are capable of inferring genomic changes, including ETS-related gene (ERG) fusions and Phosphatase and TENsin homolog (PTEN) deletion with only Hematoxylin and Eosin (H&E)-stained whole slide images, both of which are non-invasive methods by which biomarkers are predicted [7]. The survey is the only survey to incorporate both recognition and retrieval paradigms in fine-grained image analysis (FGIA), which provides more detailed perspective than prior reviews before it [6]. It also focuses on mutual methodological strategies such as discriminative descriptors choice, minimizing uncertainties, image decomposition to capture minor morphological variations [6]. The learning paradigms are compared by benchmarks evaluation based on standardized datasets, which shows how sub-category discovery is aided by retrieval techniques based on learnt feature spaces [6].

2 Fundamentals of Prostate Cancer Histopathological Analysis

The histopathological imaging is also central to the diagnosis and grading of PCa, and the Gleason scoring system is to this day the gold standard of prognostic assessment as it was first introduced in the 1960s [8]. Histopathological examination is traditionally based on visual examination of H&E stained tissue-sections by trained pathologists, which although effective, is labour and time intensive and has high inter-observer variability, especially when trying to differentiate intermediate Gleason patterns, e.g. score 7 [8]. To overcome these shortcomings, deep learning-based models have come out as potent instruments of automating and standardizing cancer grading. Several studies have demonstrated the efficacy of CNNs in classifying histopathological patches at multiple scales, leveraging architectures such as ResNet and AlexNet to achieve high accuracy in distinguishing benign from malignant regions and in fine-grained Gleason pattern recognition [9][10]. It is important to note that multi-scale analysis is shown to improve the model performance by combining contextual data present in large patch and exploiting the available numerous but smaller training samples [11]. Moreover, complex data augmentation algorithms, including image-space augmentation along with feature-space augmentation, have been revealed critical in alleviating overfitting and enhancing generalisation when limited labelled data are available [11]. In parallel, innovative imaging modalities such as MPM have been explored for label-free, subcellular-resolution imaging of unstained tissues, enabling real-time diagnostic potential when combined with deep learning frameworks [8]. Most importantly, these models do not only show a competitive classification accuracy, as reported AUC values of up to 0.943, but also produce heatmaps to visualise the spatially resolved Gleason scores, making them easier to interpret [8]. Semi-supervised methods that have active learning and uncertainty quantification have additionally alleviated the burden of annotations by ranking informative examples to be labeled [12]. Also, stain normalization and color augmentation methods have been used to enhance model sensitivity to staining procedure changes, such as those due to intraoperative frozen sections, or neoadjuvant therapy-influenced tissues [4][5]. Fully convolutional networks (FCNs) have been utilized for epithelial tissue segmentation, enabling precise localization of tumor glands, which are then assigned Gleason grades based on expert annotations—a critical step in building reliable ground truth for subsequent learning stages [12]. Deep learning models when compared to independent cohorts are found to agree with pathologists at substantially high levels as assessed by quadratic Cohen's kappa statistics (e.g., $\kappa = 0.75$) [8]. In addition, clinical nomogram incorporation of deep learning predictions, including the DL-nomogram, which involves a combination of NAFNet-generated features with clinical variables have shown improved performance against a traditional risk assessment instrument like Prostate Imaging-Reporting and Data System (PI-RADS) and Cancer of the Prostate risk Assessment (CAPRA) to predict adverse pathology and biochemical recurrence-free survival [8]. Taken together, the above developments support the transformative nature of fine-grained deep learning-based approaches to improving objectivity, reproducibility, and clinical applicability of prostate cancer histopathological analysis.

3 Fine-Grained Image Classification in Medical Imaging

3.1 Definition and Characteristics

Fine-grained learning in histopathological image analysis refers to the precise differentiation of subtle morphological variations within prostate cancer tissues, particularly in the context of Gleason grading, where inter-class differences are often minimal and intra-class heterogeneity is substantial [1]. This difficulty is even increased by the variability provided by the magnification levels which can greatly modify the feature images, even though they would be belonging to the same pathological subclass [1]. To counter these complexities, the use of deep learning models has been resorted to more and more to design hierarchical patterns of space and gain greater discriminative capacity at a microscopic scale [13]. Specifically, CNNs have shown better results in high-dimensional feature extracting whole-slide images, which allows them to be more accurately classified compared to traditional descriptors, based on texture or fractals [1]. Another area that has been studied recently is topological properties based on cellular networks, including cycles, which represent spatial relational information among nuclei and make a positive contribution to the distinction between Gleason pattern 3 and 4 tumors when combined with support vector machines [1]. Furthermore, automated systems that make use of these characteristics allow decreasing the subjectivity and improving reproducibility of pathological evaluations [1]. The incorporation of prior information like knowing that intraclass variance must be minimized and inter-subclass variance must be maximized has been demonstrated to enhance the model generalization and strength through datasets [1]. Further, DLS has proven to model Gleason patterns on a continuous scale as opposed to discrete categories, the biological continuum of tumor differentiation, and to give smooth transitional zones and uncertain areas not well agreed upon by pathologists [3][4][14]. Inter-pathologist-agreement of these fine-grained predictions is present and the predictions show gradual probabilistic gradients across grade-thresholds, indicating that DLS could quantify histological uncertainty in the same manner as clinical uncertainty would [1]. These developments highlight the potential power of fine-grained learning to address the issue of diagnostic variability and prostate cancer prognosis risk stratification, which is more leveraged.

3.2 Evaluation Metrics and Datasets

The generalization of fine-grained learning models in the analysis of prostate cancer histopathological images is largely based on standardized datasets and established metrics of performance, which are the main engine of reproducible and comparable research findings. Over the past few years, there has been a great deal of promise in using deep learning-based solutions to learn automatic Gleason grading, a task that has often remained a bottleneck of technical expertise in expertise pathologists and that is often vulnerable to inter-observer variance, especially in the intermediate range of Gleason 7 [2]. To ensure robust model assessment, researchers commonly employ metrics such as Cohen's quadratic kappa to measure agreement between automated systems and human annotators [2], while classification accuracy remains a primary indicator for multi-class fine-grained recognition tasks across subordinate categories [6]. Also, this region beneath the receiver operating characteristic curve is mainly used in binomial classification to determine between clinically significant and insignificant grades of cancer at the slide level [1]. In continuous outputs, such as the measure of the percentage of certain Gleason patterns in solid tumor areas, mean absolute error (MAE) is a sensitive predictive fidelity measure [1]. Data publically available datasets are also important to this field; an example is the TCGA dataset having detailed genomic and pathological data that can be accessed through the Genomic Data Commons portal, allowing conduct of large-scale validation studies [1]. Other cohorts, not necessarily publicly distributed because of the relevant restrictions of data usage, serve as major contributors of model training and internal validation once ethically valid [1]. Most significantly, the high-accuracy classification with up to 0.999 in recognizing cases of important prostate cancer has been made possible in benchmark data with patch-level annotations of multi scale (like that of ResNet-based architecture) [9].

These data sets are characterized by differences in their size, the level of annotation, and provision of supplementary data, e.g., hierarchical labels or survival outcomes [6]. Some can be done on an entire slide basis, others can be delivered at a patch scale or in tissue microarray (TMA) format, which is optimal to fine-style classification [9][11]. Combining such variety of resources with sophisticated deep learning structures has not only made it possible to perform with precise diagnosis but also enhanced risk rating and prediction of prognosis [8]. Additionally, systematic reviews have also emphasized the role external validation and measures of bias such as Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) with weaknesses such as overfitting and absence of external institutional validity being a frequent setback [2]. Thus, the advancements in the future will be based on the expansion of data sharing efforts and standardized check-ups that mirror actual diagnostic complexity.

4 Deep Learning Approaches for Fine-Grained Prostate Cancer Analysis

4.1 Convolutional Neural Networks and Architectures

The use of CNNs has become one of the pillars in developing fine-grained classification of images of histopathy analysis and this is more especially in prostate cancer diagnosis. The integration of deep learning architectures such as ResNet models has demonstrated exceptional performance in multi-scale patch-level analysis of digital pathology images, achieving near-perfect accuracy in identifying clinically significant prostate cancer cases [9]. It is based on hierarchical extraction of features, which can generalize fine morphological differences (representing Gleason grades), to overcome inherent subjectivity and inter-observer variability of manual assessment of pathologists [11]. In the effort to reduce the instance of model generalization under situations of limited labeled data, more sophisticated augmentation of data (including both image and feature-based augmentation) has been used and observed to be effective in improving the ability of classification to withstand [11]. Hybrid architectures combining CNNs with recurrent networks, such as long short-term memory (LSTM)-RNN applied to shearlet-transformed subbands, further exploit spatial and sequential dependencies within histopathological textures, yielding high diagnostic accuracy despite computational intensity during training [15]. Additionally, refined deep networks with regularization and batch normalization and improved hyperparameters have demonstrated better results with histopathology data being more precise and in recall measures than the state-of-the-art models [16]. New methods also consider super-resolution generative adversarial networks (SRGAN-ResNeXt) to overcome resolution constraints in low-resource imaging whose applications in visual neuroscience can aid better cell and nucleus characterisation using compressed or low-quality input [17]. Unsupervised pretraining and transfer learning Self-supervised pretraining and transfer learning paradigms compared to ImageNet-initialized and randomly initialized networks demonstrate substantial performance improvement in downstream tasks, such as classification, segmentation, and regression of histopathology data [5][14]. Also, quasi-online approaches to hard-negative mining of large patches (>112 million patches) can be used to achieve sustained model improvement to boost discriminative power by repeating inferences on large patches [1]. Besides, new generations of deep learning approaches such as NAFNet have been modified to be used not only to grade images and to predict adverse pathological outcomes and biochemical recurrence, but also to achieve higher prognostic accuracy in addition to clinical nomograms [8]. Taken together, all of these architectural developments highlight the revolutionary nature of CNNs when it comes to realizing expert-level consistency and scalability in histopathology interpretation of prostate cancer.

4.2 Attention Mechanisms and Localization

Attention incorporation into deep learning models has made breakthroughs towards the interpretation and performance of computer-aided diagnostic frameworks in pathological image examination. Specifically, the visualization methods of Gradient-weighted Class Activation Mapping (Grad-CAM) have provided healthcare professionals with the ability to evaluate the discriminative areas in histology slides of prostate cancer by highlighting areas with high activation scores and improve better model transparency [18]. The visualization techniques are useful in confirming the existence of the network in morphologically significant elements like the glandular structure or the nuclear atypia which are essential in Gleason grading [18]. Convolutional Block Attention Module (CBAM) modules have also proven to be more effective in object localization than their baseline network and have made more accurate localization of tumor areas in whole slide images (WSI) [18]. These superattention mechanisms do not only boost the accuracy of the classification, but also match model decisions to expert knowledge of pathology, and have brought the confidence of automated systems [18]. Moreover, self-attention architectures based on transformers have demonstrated effectiveness in detecting genomic alterations, including ERG fusions and PTEN deletions, directly on H&E-stained WSIs, and evidence suggests that attention processes can mediate morphological patterns with molecular phenotype [7]. Also, cellular topology derived network cycle properties have been used together with machine learning classifiers to discriminate between Gleason grade 3 and 4 tumors, which show that fine-grained discrimination is enhanced using structural connectivity patterns [18]. Although MPM and AlexNet were both able to score Gleason moderately, there is a possibility of refining region-of-interest localization in unstained tissue imaging with the use of attention maps [7]. Taken together, these strategies help highlight the relevance of spatial and semantic attention to obtain high diagnostic and clinical interpretability in the context of analyzing prostate cancer.

4.3 Weakly and Semi-Supervised Learning Methods

Weakly and semi-supervised learning paradigms have been brought out as being key solutions to the problems of lack of data annotation and inter-observer variation with prostate cancer histopathological image analysis, especially in the setting of Gleason grading [7]. The classical deep learning systems were very sensitive to pixel-based annotations which are not only time-consuming but also expensive to obtain as it takes about 900 hours of pathologists to label several datasets that cover only 115,000 mm2 of taken tissue [1]. To circumvent these limitations, recent studies have adopted multiple instance learning (MIL) frameworks that operate on slide-level labels, thereby eliminating the need for exhaustive manual segmentation while still enabling robust model training [19]. This is not just a much easier way to have considerably larger datasets that are representative of the clinical heterogeneity in the actual world, but also helps to address the reproducibility challenges which are often apparent in machine learning processes because of non-uniformity in the reporting of preprocessing procedures [20]. It is also important to note that weakly supervised models have been shown to agree with expert pathologists as well as inter-pathologist concordance rates would, resulting in Cohen's quadratic kappa values as high as 0.75 in independent test groups [7]. What is more, large-scale data exposure in such systems is of great benefit, and the performance patterns show that the performance level is expected to grow with the increase of the dataset size and quality [1]. In parallel, transformer-based hierarchical architectures have been employed to extract high-dimensional feature representations from WSIs, enabling not only accurate Gleason pattern classification but also the prediction of underlying genomic alterations—such as ERG fusions and PTEN deletions—from H&E-stained tissues alone, a task previously deemed beyond morphological assessment [7]. These developments promote the reconfigurative capabilities of weakly broadcasted methodologies in digital pathology wherein they can permit scalability, reproducibility, and clinically experienced findings without the necessity of finely grained manual markups [19].

4.4 Transfer Learning and Pretraining Strategies

Pretraining and transfer learning methods have played a significant role in the development of deep learning applications in the histopathological image analysis field (e.g., fine-grained classification tasks, like Gleason grading of prostate cancer [9]). It has been long recognized that leveraging pretrained networks helps to improve the performance of models when working with small annotated medical datasets due to the initialisation with weights obtained on natural image corpora of large-scale such as ImageNet [20, 21]. Recent research indicates that this pre training can increase the speed of convergence and generalization ability of several diagnostic tasks such as classification of cancer subtypes and prediction of survival by a significant margin [9][15]. It is noteworthy that self-supervised methods of pretraining have risen to compete with supervised ImageNet pretraining, and in classification and regression tasks, can do better, but ImageNet-based pretraining is still the frontrunner in segmentation tasks with larger architectures [21]. The combination of domain pretraining in the form of high-resolution digital pathology patches at various scales, also enhances feature representations, which are essential in aiding the detection of the subtle morphological differences in prostate cancer histology [9]. In addition, the strength of such models is bolstered by the strategies that will enable the network to automatically ignore unimportant artifacts, thus enabling better generalization to real-world clinical data, without the need to undertake data curation at a large level [19]. Experiments conducted with different sizes of the training sets demonstrate that bigger size yields higher F1 scores, but the wall increase reduction with the scale, which is significant in relation to the effectiveness of pretraining when the amount of data is limited [21]. Simultaneously, customized deep architectures—such as NAFNet and ResNet variants—have been adapted to incorporate multimodal inputs, enabling accurate prediction of adverse pathological outcomes and biochemical recurrence-free survival, outperforming conventional clinical scoring systems in AUC and risk stratification accuracy [8]. All these improvements indicate the disruptive nature of transfer learning in filling the divide between high-performance computing and clinically useful information in oncological diagnostics.

5 Emerging Trends and Advanced Methodologies

5.1 Vision Transformers in Histopathology

The use of vision transformer integration to analyze histopathological images has also been an important breakthrough in fine-grained grading of prostate cancer, using the hierarchical feature extraction qualities of the transformer-based structures to acknowledge fine morphologically-based patterns that human pathologists otherwise would have failed to notice [7]. In contrast to traditional convolutional neural networks, which exhibit great dependency on local receptive fields, vision transformers are more effective to model long-range relationships across WSIs to learn representations properly of more intricate tissue structures [2]. Recent work has shown that high accuracy of identifying clinically pertinent genomic changes including ERG fusions and PTEN deletions can be obtained using hybrid models including both transformer and CNN components using H&E-stained slides directly, and with AUC values of more than 0.85 on cross-cohort validation [7]. Nonetheless, regardless of those improvements, reproducibility is a key issue; the mismatch of data preprocessing pipelines and the lack of sufficient methodological reporting frequently worsen the situation in terms of replicating the results and thus the need to have uniform checklists to promote transparency and reliability in model development [20]. Additionally, model interpretability and clinical trustworthiness can be improved with the use of saliency maps that are based on underlying deep generative frameworks and expert-labeled regions of interest [2]. With fine-grained classification systems beyond Gleason scoring to molecular phenotyping, the interaction between transformer engine and domain knowledge is key in the creation of generalizable systems with clinically viable decision support [1].

5.2 Multi-Scale and Hierarchical Modeling

The fine-grained study of histopathological imaging is central to the proper diagnosis and prognosis of prostate cancer where morphological change capabilities should be identified by subtle morphological differences to determine the appropriate Gleason grades [19]. More recent developments in the field of deep learning have made it possible to perform multi-scale and hierarchical modeling techniques that replicate the diagnostic process of expert pathologists, which regularly examine tissue architecture at different magnifications levels to provide both cytological and architectural variations to the heterogeneity [19]. In this respect, recurrent-based CNNs have been used to estimate sequential relationships between subband features obtained using shearlet transforms in order to increase discriminability in Gleason grading problems [15]. As a response to the limitation of annotated data, the hybrid data augmentation methods that act either in both image and feature space have been crucial in improving the generalization and also helping to reduce overfitting thereby facilitating models to achieve comparable inter-observer agreement levels to that of human scientists [11]. It is particularly noteworthy that patch-level-based studies that make use of ResNet-based architecture have demonstrated near-accurate accuracy (0.999) of detecting clinically significant cancer on patch-level data, emphasizing the usefulness of deep residual learning to capture fine-gathered pathological features [9]. The net effect of these changes points to the disruptive nature of multi-scale deep learning paradigms in computational pathology, specifically in standardizing and automating Gleason grading processes.

5.3 Explainable AI and Model Interpretability

The explainable artificial intelligence (XAI) has become a crucial factor in the use of deep learning tools in analyzing histopathological images, especially in the diagnosis and grading of prostate cancer. Such a model interpretability, in addition to boosting clinical trust, also assists in regulatory approval through the provision of transparent decision-making processes [20]. Attention mechanisms and explainability maps have also been used more recently to focus attention on regions in WSIs that produce largest classification contributions; e.g. hierarchical transformer architectures have been shown to perform highly in detecting genomic alterations (e.g., ERG fusions and PTEN deletions) with the direct use of H&E-stained slides, with AUC values as high as 0.91 in validation cohorts [7]. These models produce visual explanations by visualizing the attention weight which helps pathologists compare predictions provided by the algorithm with morphological features. In the meantime, other methods less requiring human intervention which make use of saliency maps based on Deep Convolutional Generative Adversarial Networks (DCGANs), combined with expert-based semantic segmentation have further enhanced the transparency of feature extraction in Gleason grading systems [5]. Nevertheless, reproducibility is still a significant issue in the field of machine learning based histopathology, where a lack of reporting of preprocessing procedures can result in high differences in performance across implementation [20]. To solve this, the concept of standardized checklists has been suggested in order to have a full documentation of the experimental settings [20]. Moreover, network cycle characteristics of cellular spatial arrays provide an understandable representation, that is, they capture topological motifs of glandular patterns and are more effective than traditional texture- and fractal-based measures to identify Gleason grade 3 and grade 4 tumors [5]. These techniques collectively highlight how crucial it is to build interpretability into the design of a model instead of viewing it as an ex post hoc offer.

The intersected viewpoints of various operationalization of interpretability paradigms, spanning attention-oriented visualization [7], saliency-directed analysis [5], topological feature design [5], to localization-centered heatmap [20], demonstrates the multidimensional stepping in multiple directions of trustful AI in digital pathology. It is interesting that high accuracy rates (e.g. 98% ± 2% in Gleason grading [5]) indicate that these systems are technically feasible, but whether they will be adopted by clinicians depends on whether they can give rationales that humans can understand. Transformer-based models have demonstrated specific potential in the generation of spatially resolved explainability maps, which are consistent with immunohistochemical validation, particularly in subclonal PTEN loss samples in which the area predicted by the model correlates with observed ground truth ($r = 0.58$, $P = .0097$) [7]. Additionally, non-invasive imaging data uses such as MPM can decrease the utilization of standard staining protocols, which provides a prospect of intraoperative diagnostics which could be backed by interpretable deep learning architectures with AUCs greater than 0.94 [20]. However, reproducibility may only be achieved through stringent standardization of data curation pipelines, which is raised in independent reproduction of state-of-the-art algorithms where unspecified preprocessing operations resulted in significant decreases in performance [20]. As such, future evolution needs to find a compromise between architectural sophistication and methodological openness by inculcating some level of interpretability in all the layers of computational oncology, which includes the stage of feature extraction and the final prediction.

6 Conclusion

Deep learning-based fine-grained learning has brought a revolutionary phase of automated interpretation of PCa histopathology as a major alternative to the subjectiveness nature of classical Gleason grading and its work-intensive nature. This survey has systematically highlighted the evolution from conventional visual assessment to sophisticated CAD systems. Core to this progress is the application of diverse architectures, notably CNNs and emerging Vision Transformers, which capture the subtle, minute morphological differences that define fine-grained classification.

The successful implementation of multi-scale and hierarchical modeling to simulate the whole workflow of a pathologist as part of its key contributions and classification accuracies that can compete with the inter-pathologist agreement. Importantly, the discipline is no longer just detecting but based on an excavation of the pathology. Such tools as the attention mechanisms and the XAI do not just enhance performance, but are allowing clinical trust in that the discriminative areas can be localized, and underlying genomic changes (e.g., ERG fusions) can be inferred directly based on the H&E-stained images. Moreover, the adoption of weakly and semi-supervised learning paradigms, such as MIL, is effectively addressing the pervasive challenge of data annotation scarcity, allowing for the deployment of models on large, real-world datasets.

Finally, these deep learning networks with fine grains are optimizing the prognostic process providing subtle, probabilistic forecasts, which measure histological ambiguity. Although this field still has challenges especially in terms of reproducibility owing to differences in preprocessing and the necessity of standardized reporting, the future of this field is toward the development of highly precise, scaleable, and interpretable tools. It can be anticipated that the interaction of these sophisticated systems with the clinical practice can result in adopting PCa risk stratification that is standard, diagnostic variation decrease, and prognostic precision.

References

1. Nagpal, K., Foote, D., Liu, Y., Chen, P. H. C., Wulczyn, E., Tan, F., ... & Stumpe, M. C.: Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer. NPJ digital medicine 2(1), 48 (2019)
2. Rabilloud, N., Allaume, P., Acosta, O., De Crevoisier, R., Bourgade, R., Loussouarn, D., Rioux-Leclercq, N., Khene, Z.-E., Mathieu, R., Bensalah, K., Pecot, T., & Kammerer-Jacquet, S.-F.: Deep learning methodologies applied to digital pathology in prostate cancer: A systematic review. Diagnostics 13(16), 2676 (2023)
3. Li, L., Pan, X., Yang, H., Liu, Z., He, Y., Li, Z., ... & Zhang, L.: Multi-task deep learning for fine-grained classification and grading in breast cancer histopathological images. Multimedia Tools and Applications 79(21), 14509-14528 (2020)
4. Epstein, J. I., Egevad, L., Amin, M. B., Delahunt, B., Srigley, J. R., Humphrey, P. A., & Grading Committee.: The 2014 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma: definition of grading

patterns and proposal for a new grading system. The American journal of surgical pathology 40(2), 244-252 (2016)

5. Khurd, P., Grady, L., Kamen, A., Gibbs-Strauss, S., Genega, E. M., & Frangioni, J. V.: Network cycle features: Application to computer-aided Gleason grading of prostate cancer histopathological images. In: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 1632–1636. IEEE, Chicago, IL, USA (2011)
6. Wei, X.-S., Luo, J.-H., Wu, J., & Zhou, Z.-H.: Fine-grained image analysis with deep learning: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence 44(12), 9477–9501 (2022)
7. Erak, E., Oliveira, L. D., Mendes, A. A., Dairo, O., Ertunc, O., Kulac, I., ... & Lotan, T. L.: Predicting prostate cancer molecular subtype with deep learning on histopathologic images. Modern Pathology 36(10), 100247 (2023)
8. Gu, W.-J., Liu, Z., Yang, Y.-J., Zhang, X.-Z., Chen, L.-Y., Wan, F.-N., Liu, X.-H., Chen, Z.-Z., Kong, Y.-Y., & Dai, B.: A deep learning model, NAFNet, predicts adverse pathology and recurrence in prostate cancer using MRIs. npj Precision Oncology 7(1), 134 (2023)
9. Kondejkar, T., Al-Heejawi, S. M. A., Breggia, A., Ahmad, B., Christman, R., Ryan, S. T., & Amal, S.: Multi-scale digital pathology patch-level prostate cancer grading using deep learning: use case evaluation of DiagSet dataset. Bioengineering 11(6), 624 (2024)
10. Arvaniti, E., Fricker, K. S., Moret, M., Rupp, N., Hermanns, T., Fankhauser, C., ... & Claassen, M.: Automated Gleason grading of prostate cancer tissue microarrays via deep learning. Scientific Reports 8(1), 12054 (2018)
11. Karimi, D., Nir, G., Fazli, L., Black, P. C., Goldenberg, L., & Salcudean, S. E.: Deep learning-based Gleason grading of prostate cancer from histopathology images—role of multiscale decision aggregation and data augmentation. IEEE Journal of Biomedical and Health Informatics 24(5), 1413–1426 (2019)
12. Singhal, N., Chattopadhyay, N., Samanta, P., & Bonthu, S.: Prostate Cancer Diagnosis and Grading in Whole Slide Images of Core Needle Biopsies. Scientific Reports 12(12), 3383 (2022)
13. Gurumoorthy, R., & Kamarasan, M.: Wildebeest habit optimizer with deep learning-based histopathological image analysis for breast cancer diagnosis. International Journal of Engineering Trends and Technology 71(7), 233–243 (2023)
14. Palmieri, C., & Grieco, V.: Proposal of Gleason-like grading system of canine prostate carcinoma in veterinary pathology practice. Research in Veterinary Science 103, 11–15 (2015)
15. Uthappa, P. P., & Sharma, S. L.: Hybrid Unified Deep Learning Network for Highly Precise Gleason Grading of Prostate Cancer. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 899-903. IEEE, Berlin, Germany (2019)
16. El-Ghany, S. A., Azad, M., & Elmogy, M.: Robustness fine-tuning deep learning model for cancers diagnosis based on histopathology image analysis. Diagnostics 13(4), 699 (2023)
17. Juhong, A., Li, B., Yao, C. Y., Yang, C. W., Agnew, D. W., Lei, Y. L., ... & Qiu, Z.: Super-resolution and segmentation deep learning for breast cancer histopathology image analysis. Biomedical Optics Express 14(1), 18–36 (2022)
18. Woo, S., Park, J., Lee, J.-Y., & Kweon, I. S.: CBAM: Convolutional Block Attention Module. Lecture Notes in Computer Science (ECCV) 11211, 3–19 (2018)
19. Campanella, G., Hanna, M. G., Geneslaw, L., Miraflor, A., Werneck Krauss Silva, V., Busam, K., Brogi, E., Reuter, V. E., Klimstra, D. S., & Fuchs, T. J.: Clinical-grade

computational pathology using weakly supervised deep learning on whole slide images. Nature Medicine 25(8), 1301–1309 (2019)
20. Fell, C., Mohammadi, M., Morrison, D., Arandjelovic, O., Caie, P., & Harris-Birtill, D.: Reproducibility of deep learning in digital pathology whole slide image analysis. PLoS Digital Health 1(12), e0000145 (2022)
21. Ciga, O., Xu, T., & Martel, A. L.: Self supervised contrastive learning for digital histopathology. Machine Learning with Applications 7, 100198 (2022)