Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Target Device Specific Optimized Deep Learning Model For Oral Cancer Diagnosis

Aarti Yadav[1] , Dr. Surendra Yadav[1] , Dr. Navneet Garg[2]

[1]*Department of Computer Science Engineering Vivekananda global university, Jaipur (Raj), India*
raoaarti0@gmail.com, surendra.yadav@vgu.ac.in
[2]*MD, Radiology Shubh Plus Diagnostics, Jaipur (Raj), India*
shubhplusdiagnostics@gmail.com

This study presents a novel methodology aimed at enhancing the early detection of Oral Squamous Cell Carcinoma (OSCC) through the application of deep learning methodologies on histopathological samples. We delve into the intricacies of optimizing the performance of two widely recognized convolutional neural network architectures, namely DenseNet-169 ,VGG-19 and ResNet, with the primary objective of bolstering both the accuracy and faster inferencing of OSCC identification processes. Utilizing the advanced capabilities provided by the OpenVINO toolkit, we embark on a comprehensive exploration through benchmarking experiments across a spectrum of hardware configurations, encompassing variations such as CPU, GPU, and hybrid setups. This rigorous evaluation enables us to scrutinize the inference performance of the models under diverse computational environments. Through meticulous analysis, we discern and delineate optimal deployment strategies tailored for real-world application scenarios. In this endeavor, we strike a delicate balance between computational resource utilization and inference accuracy, thereby furnishing valuable insights into the realm of deep learning-based solutions for advancing the early detection of OSCC. This culminates the huge potential lying ahead in using deep learning methodologies to drive some major strategies for the enrichment of diagnostic protocols in OSCC and sets the stage for transformative impacts in clinical practice and medical research.

Keywords- Deep learning, Oral Squamous Cell Carcinoma, Histopathological samples, DenseNet, VGG,ResNet Performance optimization, OpenVINO, Benchmarking.

I. INTRODUCTION

OSCC has been one of the most common forms of malignancy in the world and, together with its highly aggressive diagnosis usually at a late stage, it poses a serious health problem. The capability to detect OSCC at an early stage of the disease remains related to improved prognosis and a reduced mortality rate associated with this disease, even with comprehensive advancements in medical imaging and diagnostic methodologies [1]. Interest in using new technologies, with more deep learning methodologies, is growing for researching and developing better screening and diagnostic tools related to cancer detection across different medical domains.

Deep learning has of late attracted a great deal of research interest due to its great potential for making a sea change in medical image analysis and pathology interpretation. Researchers have been relatively successful with the automatization of detection and classification of various cancer types, including OSCC, by training CNN models on large datasets of histopathological samples. Such deep learning models would, therefore, be useful in supplementing the conventional diagnostic model for more accurate and faster identification of malignant lesions if suitably optimized and deployed.

In this backdrop, the present paper proposes a new intervention with an objective to leverage deep learning methodologies for early diagnosis of OSCC through histopathological sample analysis. Specifically, our research focuses on the optimization of performance of two of the more widely known CNN architectures: DenseNet-169, VGG-19, and ResNet [3], renegade for image classification, which are applied in this paper. These models are further fine-tuned, and their applicability to OSCC detection is considered with the purpose of increasing the accuracy and efficiency of diagnostic processes [4].

Concretely, our approach uses the Open Visual Inference and Neural Network Optimization toolkit, OpenVINO—an end-to-end complete suite of tools and libraries for fast development and easy deployment of a wide span of deep learning models on any hardware. To estimate the inference performance of the optimized models on various hardware configurations, a large collection of rigorous benchmarking experiments is maintained. Such careful evaluation will be very instrumental in determining an optimal deployment strategy in any real-world clinical setting where computational resources and inference accuracy are most crucial [5].

Our paper helps further the efforts toward innovative cancer diagnostics for improved patient outcomes by explaining the prowess and limitations of deep learning-based OSCC detection methodologies. This in turn assists in bringing state-of-the-art technology into entrenched existing medical practices that help reduce the burden of OSCC, which would be much aided by early detection and timely intervention. Eventually, this can result in changes of clinical practice itself—by using cancer screening with effective protocols and treatment strategies that are personalized [5].

In the context of oral squamous cell carcinoma presenting a formidable challenge to healthcare in India, our work assumes great significance. OSCC continues to be among the major causes of morbidity and mortality due to cancer in the country, mainly related to tobacco and betel nut consumption. Despite advances in medical technology, late diagnosis of OSCC still occurs and prevents the best possible patient outcome and survival. Among the various compelling needs to establish innovative strategies for enhancing diagnosis is the fact that early detection and accurate staging represent an urgent necessity in regard to OSCC, as it enables its early intervention and improves prognosis [5].

Objectives of Research:

The present study mainly aims at developing and testing a computer-aided diagnostic system to aid in the early detection of OSCC using deep learning and image analysis. Improvement in sensitivity and specificity would allow for timely intervention, hence better health outcomes for the patients. It is expected that observer bias in traditional pathology will be reduced by these automated techniques of analysis, leading to a better consistency of diagnosis. The system performance will also be optimized with regard to computational efficiency and speed of inference so that results can be returned rapidly and reliably, suitable for a clinical application. Feasibility and efficacy of integration into routine clinical practice will be validated by the evaluation of the system for impact on precision, efficiency, and patient outcomes against diagnostic benchmarks established. Conclusively, the study aims to share its conclusions with the medical fraternity by publishing and presenting. This will mean global development of OSCC diagnosis and management for the betterment of patient care.

II. HISTOPATHOLOGY ROLE IN CANCER ANALYSIS

Oral cancer diagnosis and management, particularly oral squamous cell carcinoma (OSCC), is of the essence in India because of its high prevalence and serious health consequences. OSCC accounts for a significant percentage of the nation's burden of cancer and is mainly caused by various habits such as tobacco and betel nut consumption [5]. Their diagnosis in most cases remains undistinguished until the progression of the disease to advanced stages, which clinically presents with poor survival rates [5]. Therefore, in an attempt to enhance diagnosis and decrease observer bias, computer-based systems have been explored to help pathologists in identifying and assessing malignancies. Diagnosis and staging of oral cancer are important for its early initiation of proper treatment and better prognosis because the intensity and progression of the disease at its different locations and size are variable [5].

Affected human tissue for diagnosis is obtained by collecting patients tissue samples (biopsies) during a clinical examination and then sent to pathology laboratories for analysis. These samples are usually stained with Hematoxylin and Eosin (H&E) for the identification of tissue structures and are viewed under the microscope by pathologists [6]. Oral cancer, principally oral squamous cell carcinoma, is a serious health problem in India due to the addiction to tobacco consumption, which causes high incidence and late diagnosis. Computer-aided techniques can become one of the most desired ways to increase diagnosis accuracy and speed, enabling pathologists to spend more time on key cases and probably improving outcomes for this very common and complex health problem [6].

III. LITERATURE SURVEY

Warin et al. (2022) found DenseNet-169 to be the best model, with an AUC of 1.00 for OSCC and 0.98 for OPMDs. Consequently, it can be said that the potential of CNN models in this task is very high and promises much for the early detection of oral cancer, beyond general practitioners alone.

Mentel et al. (2021) investigated breath analysis as a medium for the diagnosis of OSCC, establishing different compound signatures in breath samples from OSCC patients compared to healthy ones. They achieved accuracies between 86% and 90% for the classification of healthy versus patient breath by means of machine learning. This proof-of-concept study indicates that further evaluation and optimization are required.

Alabi et al. (2021) explained the role of deep machine learning in early OSCC detection, citing innovations made to date on the analysis of medical imaging. Their study identifies deep learning as playing a very critical role in advancing precision medicine for OSCC, spanning detection, classification, segmentation, to synthesis [8].

Musulin et al. (2021) used AI-assisted technologies to bring in more objectivity and accuracy in grading by analyzing histopathology images for OSCC. They compared several deep learning techniques that could use AI's potential in analyzing intricated texts and structures of oral cancer tissues [9].

Jubair et al. (2022) investigated a lightweight CNN model for EfficientNet-B0 that returned an accuracy of 85.0%, specificity of 84.5%, sensitivity of 86.7%, and an AUC of 0.928. Thus, deep CNNs can be used with very cheap embedded vision devices for oral cancer diagnosis in low resource settings, significantly improving quality and ability to screen [10].

Rahman et al. (2022) presented the severity of oral cancer as a very common and life-threatening disease, with a high mortality rate since it is the most common cancer in the world, responsible annually for more than 300,335 deaths. Though biopsy is a common diagnostic approach, often microscopic examination remains inadequate and is liable to human error [11].

IV. MODEL OPTIMIZATION

Diagnostic difficulties in oral squamous cell carcinoma lead to the major problem of late-stage detection with poor patient outcomes. Therefore, this study focuses on developing a computer-aided diagnostic system that integrates deep learning and image analysis techniques for identifying OSCC at an early stage. Clinical efficiency is foreseen to increase, hence increasing quality care to patients through decreased observer bias and increased diagnostic accuracy for this system. This shall be through data collection, preprocessing, model training, and optimization. The results shall be shared for the benefit of global practices in OSCC diagnosis and management. Below is the methodology that has been followed up for the proposed approach:

- Data Collection: Begin with the collection of a full dataset of histopathological samples of oral squamous cell carcinoma. These will be drawn from pathology labs and medical institutions; hence, these specimens will comprise all kinds of patients, based on demographic factors, tumor characteristics, and disease stages.
- Data Preprocessing: Histopathological images preprocessing maintains a standard view, resolution, and color space. Depending on requirements, normalization may require image resizing through windowing, noise reduction, or blend shots to be applied for improvement in quality and homogeneity in the dataset [12].
- Model Choice: Deep learning models relevant to theOSC detection task are chosen in this phase. Considering their performance in image classification tasks along with compatibility with the histopathological data, DenseNet-169 and VGG-19 have been selected as primary models.
- Model Training: Use the pre-processed histopathological images to train the models of choice. Herein, the dataset will be divided into three sets, namely, the training, validation, and testing sets. Transfer learning techniques will not only be applied to speed up the training/training towards better convergent trained models but also for better convergence.
- Testing: The trained models, later on, are tested in the test dataset, calculating accuracy, sensitivity, specificity, and other performance measures. Outputs from model performances

can be further visualized in a receiver operating characteristic graph or a confusion matrix to understand where improvements occur.

- Optimization: At this stage, the performance of trained models is optimized through a host of strategies, with perhaps the most notable being hyperparameter tuning, change of architecture, and data augmentation. Other methods include gradient clipping, learning rate scheduling, and regularization, a group of strategies to guarantee, among others, generalization by preventing models from overfitting [].
- Hardware Acceleration: A further path of combination with hardware acceleration frameworks like Intel's OpenVINO toolkit could be used to help accelerate inference and deployment. This approach will further enhance the execution speed for a deep learning model over many hardware platforms, even just on any CPU, GPU, or other specialized accelerator.
- Clinical validation: Clinical validity of the optimized models will be checked on diagnostic gold-standard protocols and assessments independently by expert pathologists. Simulated scenarios assess for feasibility and effectiveness in real-world deployment integrated in routine clinical practice.
- Iterative Refinement to the Component: The methodology undergoes iterative refinement by klinische validation and performance evaluation. The model's parameters, training data, and optimization techniques get iteratively changed to build a robust diagnostic system.
- Final Documentation and reporting—It will document a detailed form of methodology, results, and findings. Additionally, it would disseminate research outcomes in peer-reviewed publications, conference presentations, and knowledge-sharing.

To this respect, model optimization becomes very critical. Given that the risk involved is a diagnosis concerning cancer, the performance of the deep learning model w.r.t. reliability and accuracy should be maximized. Therefore, methods for model capacity improvement, in terms of improved discernment of more subtle patterns indicative of early-stage OSCC, include hyperparameter tuning, refinement in model architecture, or other advanced training strategies. It will also reduce false positives and false negatives in the phase of model optimization, thus increasing the reliability of the diagnostic tool. Further, with thorough fine-tuning of the model, it will generalize well across a wide variety of datasets and imaging conditions for their eventual variegation into a select system to assist clinicians for rendering more accurate diagnoses and timely interventional decisions.

V. RESULT ANALYSIS

5.1 VGG-19 Model

A comprehensive benchmarking analysis aimed at evaluating the performance of the VGG-19 convolutional neural network (CNN) architecture on different hardware configurations. In the realm of deep learning, optimizing the performance of models is paramount, particularly in medical applications where timely and accurate diagnosis can significantly impact patient outcomes. This section focuses on assessing the inference speed and efficiency of the VGG-

19 model when deployed on various hardware devices, including CPU, GPU, and hybrid setups, using the OpenVINO toolkit [14].

Deep learning models, such as VGG-19, have demonstrated remarkable capabilities in image classification tasks, making them valuable tools for medical image analysis. One such critical application is the early detection of diseases like oral squamous cell carcinoma (OSCC), where timely diagnosis is essential for effective treatment and improved prognosis. By benchmarking the performance of the VGG-19 model on different hardware configurations, we aim to identify optimal deployment strategies that balance computational resources with inference accuracy, ultimately enhancing the efficiency of OSCC detection and other medical imaging tasks.

This section demonstrates benchmarking experiments conducted on different hardware configurations using a deep learning model (VGG-19) for asynchronous inference.

CPU Benchmarking: The first benchmarking experiment is performed on a CPU (Intel Core i7-8750H CPU @ 2.20GHz). The model is evaluated for 15 seconds using asynchronous inference, and the results indicate an average latency of 358.92 milliseconds (ms) and a throughput of 11.13 frames per second (FPS). The experiment comprises 172 iterations.

AUTO Benchmarking: The second experiment employs the "AUTO" setting, which allows the system to automatically select the hardware configuration based on available devices. The benchmarking duration and setup remain consistent with the CPU benchmark. Here, the average latency is reduced to 237.81 ms, with an increased throughput of 16.73 FPS over 256 iterations.

GPU Benchmarking: The third experiment utilizes a GPU (Intel Graphics [0x3e9b] - integrated GPU). Compared to the CPU benchmark, the GPU-based inference demonstrates significantly lower latency, with an average of 90.50 ms, and a higher throughput of 22.03 FPS. This experiment involves 334 iterations.

MULTI:CPU,GPU Benchmarking: The final experiment involves a hybrid configuration utilizing both CPU and GPU (MULTI:CPU,GPU). This setup aims to leverage the combined processing power of both devices. The results show an average latency of 90.50 ms and a throughput of 26.01 FPS across 402 iterations, indicating improved performance compared to individual CPU and GPU setups.

Overall, these benchmarking experiments provide insights into the performance of the VGG-19 model on different hardware configurations, enabling informed decisions regarding model deployment and optimization for real-world applications.

Table 2. VGG-19 Latency Evaluation

Hardware Configuration	Model	DeviceCount	Duration (ms)	Median Latency (ms)	Average Latency (ms)	Minimum Latency (ms)
VGG-19	CPU	17	2154	55.53	351.28	344.20
VGG-19	AUTO	256	15300	50.18	237.81	197.24
VGG-19	GPU	334	15158	90.45	90.50	45.21
VGG-19	CPU+GPU	402	15454	None	None	26.01

Table 3. VGG-19 Throughput

Hardware Configuration	Model	Iteration Count	Duration (ms)	Throughput (FPS)	CPU
VGG-19	CPU	17	2154	11.13	AUTO
VGG-19	AUTO	256	15300	16.73	AUTO
VGG-19	GPU	334	15158	22.03	AUTO
VGG-19	CPU+GPU	402	15454	None	MULTI:CPU,GPU

Evaluation for table 2:

VGG-19, CPU Configuration: This configuration shows medium performance—relatively low device count and duration—but, compared to other configurations, has higher median and average latencies, thus indicating slower inference times.

AUTO Configuration (VGG-19, AUTO): Compared to the configuration over CPU, the AUTO configuration comes with a higher number of devices and longer durations. It improved the median and average latencies, which, in turn, is faster in inference time, though it has large variation from minimum to maximum latencies.

VGG-19, GPU: Compared with the AUTO configuration, it has more devices used with a longer duration. The median and average latencies are lower than a CPU configuration, which means that inference times are faster and the performance is better.

Testing CPU+GPU Configuration—VGG-19, CPU+GPU: The hybrid setup is a combination of both the resources of a central processing unit and a graphics processing unit. The results include the highest count of devices and duration in this case. Although latency statistics show up as "None" in the table, throughput is relatively high compared to the other scenarios, thus showing efficiency in performance.

Overall, the configuration seems to provide one of the best trade-offs between the number of devices, duration, and latency, thus indicating good and efficient performance of the VGG-19 model. However, further details on latency statistics for the CPU+GPU configuration would provide a clearer evaluation of its performance.

Table 3 demonstrates the performance of the VGG-19 model across various hardware configurations by evaluating key metrics such as iteration count, duration, and throughput in frames per second (FPS). When running on a CPU, VGG-19 achieved a throughput of 11.13 FPS with 17 iterations completed in 2154 milliseconds. In an automatic (AUTO) configuration, likely involving automatic hardware selection, the model processed 256 iterations in 15300 milliseconds with a throughput of 16.73 FPS. The GPU configuration significantly enhanced performance, completing 334 iterations in 15158 milliseconds with a throughput of 22.03 FPS. The combined CPU and GPU setup processed the highest number of iterations, 402 in 15454 milliseconds, although the throughput was not provided. This comparison highlights the substantial performance benefits of utilizing GPUs for deep learning models, showcasing their ability to process more frames per second compared to CPUs, and emphasizing the importance of hardware optimization in improving model efficiency.

5.2 DenseNet-169

The overall concept of the benchmarking results is to evaluate the performance of the DenseNet-169 model across different hardware configurations and inference modes. The aim is to identify the most efficient setup for conducting inference tasks with this deep learning model.

The benchmarking is conducted using four different hardware configurations as shown in Table 4:

- CPU: Utilizing the CPU for inference.
- AUTO: Automatically selecting the hardware configuration.
- GPU: Employing the GPU for inference.
- MULTI:CPU,GPU: Simultaneously utilizing both the CPU and GPU for inference.

Benchmarking with CPU:

- **Count**: 436 iterations
- **Duration**: 15085.43 ms
- **Latency**:
    - Median: 131.00 ms
    - AVG: 137.99 ms

    - MIN: 72.00 ms
    - MAX: 214.65 ms
- **Throughput**: 28.90 FPS

Explanation: When running inference on the CPU, the model processed 436 iterations in 15 seconds. The latency statistics indicate that the median latency (time taken for each inference) is 131.00 ms, with an average latency of 137.99 ms. The throughput, which represents the number of frames processed per second, is 28.90 FPS.

Benchmarking with AUTO Configuration:

- **Count**: 252 iterations
- **Duration**: 15398.84 ms
- **Latency**:
    - Median: 241.55 ms
    - AVG: 242.91 ms
    - MIN: 62.90 ms
    - MAX: 444.62 ms
- **Throughput**: 16.36 FPS

Explanation: The AUTO configuration selects the hardware automatically. In this case, it might have selected the CPU or GPU based on system settings. The throughput is lower compared to the CPU configuration, indicating that the selected hardware may not be optimized for this task.

Benchmarking with GPU:

Count: 412 iterations

Duration: 15126.81 ms

Latency:
Median: 72.65 ms
AVG: 73.19 ms
MIN: 38.77 ms
MAX: 90.02 ms
Throughput: 27.24 FPS

Explanation: Running inference on the GPU results in lower latencies compared to the CPU configuration. The GPU configuration processed more iterations with faster median and average latencies, leading to a higher throughput of 27.24 FPS.

Benchmarking with MULTI:CPU,GPU Configuration:
Count: 624 iterations
Duration: 15183.84 ms
Throughput: 41.10 FPS

Explanation: This configuration utilizes both the CPU and GPU simultaneously as shown in Table 5, resulting in the highest throughput of 41.10 FPS. While detailed latency statistics are not provided, the combined processing power of the CPU and GPU leads to faster inference times and higher throughput.

Table 4. VGG-19 Latency Evaluation

Hardware Configuration	Model	Device Count	Duration (ms)	Median Latency (ms)	Average Latency (ms)	Minimum Latency (ms)	Maximum Latency (ms)
CPU	DenseNet-169	17	15085.43	131.00	137.99	72.00	214.65
AUTO	DenseNet-169	252	15398.84	241.55	242.91	62.90	444.62
GPU	DenseNet-169	412	15126.81	72.65	73.19	38.77	90.02
MULTI:CPU,GPU	DenseNet-169	624	15183.84	-	-	-	-

Table 5. VGG-19 Throughput Evaluation

Hardware Configuration	Model	Device Count	Throughput (FPS)
CPU	DenseNet-169	17	28.90
AUTO	DenseNet-169	252	16.36
GPU	DenseNet-169	412	27.24
MULTI:CPU,GPU	DenseNet-169	624	41.10

Each configuration is evaluated based on several performance metrics, including duration (the time taken for the benchmarking process), latency (the time taken for each inference), and throughput (the number of inferences processed per second).

The results show that:

Utilizing the GPU generally leads to lower latencies and higher throughputs compared to using the CPU alone. The MULTI:CPU,GPU configuration, which leverages both the CPU and GPU simultaneously, achieves the highest throughput, indicating superior performance. Conversely, the AUTO configuration may not always select the most optimized hardware for the task, resulting in lower throughput compared to manually selecting the GPU.

These results generalize that GPU acceleration and a multi-device configuration significantly improve the performance of DenseNet-169 for inference tasks. Researchers and developers can choose appropriate hardware configurations to fine-tune their deep learning inference, aiming to improve efficiency and speed if not to bring better performance in model deployment in real-world applications.

Comparison between VGG-19 and DenseNet-169:

Latency Comparison:

Latency varies for both VGG-19 and DenseNet-169 across hardware configurations. In general, DenseNet-169 has a lower median latency in comparison with VGG-19, particularly on the GPU configuration options.

Throughput Comparison:

In most cases, DenseNet-169 throughputs are better than those of VGG-19 on similar hardware configurations. DenseNet-169 achieves higher frames per second (FPS) values, indicating superior performance in processing frames per second.

Overall, DenseNet-169 shows competitive performance compared to VGG-19, with lower latency and higher throughput in most cases. However, the choice between the two models may depend on specific application requirements and hardware constraints.

5.3 ResNet

ResNet, short for Residual Neural Network, is a deep learning architecture that was introduced to address the problem of vanishing gradients in very deep neural networks. It was proposed by Kaiming He, et al. in their paper "Deep Residual Learning for Image Recognition," which won the Best Paper Award at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Here's an overview of the ResNet model:

Before the advent of ResNet, deep neural networks struggled with vanishing gradients, where gradients became very small during backpropagation, making training difficult and limiting the effective depth of networks. ResNet revolutionized this by introducing residual learning, which uses skip connections to bypass one or more layers. In this approach, residual blocks with shortcut connections allow the network to learn residual functions relative to identity mappings, simplifying the optimization of deep networks. The architecture of ResNet consists of these residual blocks, typically featuring two or three convolutional layers with batch normalization and ReLU activations, and is available in variants like ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152, named for their layer counts. The primary advantage of ResNet is that its skip connections facilitate direct information flow, mitigating the vanishing gradient issue and enabling the training of very deep networks. Consequently, ResNet has achieved state-of-the-art performance in various computer vision tasks, including image classification, object detection, and image segmentation, and is widely used in applications such as medical image analysis and autonomous driving.

In simple words, ResNet distorted deep learning by the introduction of skip connections and residual learning; it enables the training of very deep neural networks with improved efficiency and performance. It has far-flung impacts on many applications of computer vision and remains to this day one of the cornerstones within the field of deep learning.

Result Analysis:

Table 6. ResNet Throughput Evaluation

Hardware Configuration	**Frame Rate (FPS)**
GPU	22.84
Multi-device	38.6
CPU	18.65

Performance Analysis of ResNet on Different Hardware Configurations

Table 6: Evaluation : GPU—Graphics Processing Unit: The ResNet model achieved a frame rate of 22.84 FPS when running on a GPU. This means that nearly 22.84 frames are being processed every second through the ResNet model with a GPU. Very importantly, the GPU has a parallel architecture that makes them eminently suitable for deep learning tasks, which typically involve large amounts of matrix multiplications and convolutions while computing neural networks. This frame rate does, however, connote that actual performance with this ResNet model will have strong dependence on the specific model of the GPU and specifications in place, and the degree of optimization..

Multi-device Setup: In a multi-device configuration, the ResNet model achieves a higher frame rate of 38.6 FPS. This significant improvement indicates that leveraging multiple devices, such as multiple GPUs or a combination of GPUs and CPUs, can substantially enhance processing speed compared to a single GPU. Multi-device setups are common in large-scale deep learning applications, where distributing the workload across multiple devices accelerates the training or inference process. The higher frame rate compared to a single GPU suggests effective resource utilization and efficient parallelization of computations.

Central Processing Unit (CPU): Running on a CPU alone, this ResNet model gives 18.65 FPS. Generally, a CPU is not as suitable for deep learning as a GPU, since it is oriented more towards sequential processing. Although modern CPUs are capable of running deep learning inference, they usually have slower processing speeds compared to the same models run on a GPU. The lower frame rate reflects that the CPU is not as efficient at processing the ResNet model in comparison with processing using the GPU. However, the values of the CPUs are not completely lost in applications where huge parallelism of GPUs is simply not required or when resources on the GPU are limited. From here, one can see clearly how hardware configuration reaches performance effects on deep models. When applying parallel architecture consideration, it is suitably seen that the use of GPUs is better in the performance of deep learning tasks, and multiple devices can be used to improve processing even further.

Although generally slower for deep learning, CPUs still have a role in some scenarios and tasks.

VI. CONCLUSION AND FUTURE WORK

A. Conclusion

In this paper, we have used deep learning techniques to detect Oral Squamous Cell Carcinoma at its initial stage from given histopathological samples. We have been able to increase DenseNet-169, VGG-19, and ResNet architectures to get better accuracy and effectiveness for the purpose of clinical diagnosis. We have shown the results of benchmarking on different hardware configurations, such as CPU, GPU, and hybrid setups, to determine the best deployment strategy that allows for computational resources to be balanced against inference accuracy.

Results show that VGG-19, while being run on multi-devices (CPU+GPU) in parallel mode, produces maximum model throughput of 41.10 FPS. Our findings underline the potentials of deep learning-based methods, which enable the improvement of early OSCC detection by deriving valuable insights for clinical practice and medical research.. The integration of computer-aided diagnostic systems into routine clinical workflows holds promise for reducing observer bias and improving diagnostic consistency, ultimately leading to better patient outcomes. Oncological features identified in Class Activation Maps are verified by medical expert for presence of pleomorphism, cellular atypia, increased mitotic activity, cohesiveness and vascularity in investigated samples of Oral Cancer.

B. Future Work

Despite the promising results achieved in this study, several avenues for future research and development remain. Investigating more advanced neural network architectures tailored for histopathological image analysis could further improve diagnostic accuracy and efficiency. Exploring data augmentation and transfer learning techniques to leverage larger datasets may enhance the models' generalization capabilities, especially in diverse clinical settings. Optimizing model inference for real-time deployment on edge devices could enable point-of-care diagnosis and telemedicine applications. Collaborating with healthcare institutions to seamlessly integrate computer-aided diagnostic systems into existing clinical workflows is crucial for ensuring user-friendliness and regulatory compliance. Additionally, conducting rigorous validation studies and clinical trials is essential to evaluate the performance and impact of the proposed systems on patient outcomes in real-world scenarios. Addressing these areas can further advance computer-aided diagnosis for OSCC, contributing to improved early detection, treatment, and patient survival rates.

REFERENCES

[1] Basu, K., Sinha, R., Ong, A., & Basu, T. (2020). Artificial intelligence: How is it changing medical sciences and its future?. Indian journal of dermatology, 65(5), 365.

[2] Bishnoi, L., & Singh, S. N. (2018, January). Artificial intelligence techniques used in medical sciences: a review. In 2018 8th international conference on cloud computing, data science & engineering (Confluence) (pp. 1-8). IEEE.

[3] Guan, J. (2019). Artificial intelligence in healthcare and medicine: promises, ethical challenges and governance. Chinese Medical Sciences Journal, 34(2), 76-83.

[4] Winkler-Schwartz, A., Bissonnette, V., Mirchi, N., Ponnudurai, N., Yilmaz, R., Ledwos, N., ... & Del Maestro, R. F. (2019). Artificial intelligence in medical education: best practices using machine learning to assess surgical expertise in virtual reality simulation. Journal of surgical education, 76(6), 1681-1690.

[5] Chai, A. W. Y., Lim, K. P., & Cheong, S. C. (2020, April). Translational genomics and recent advances in oral squamous cell carcinoma. In Seminars in cancer biology (Vol. 61, pp. 71-83). Academic Press.

[6] Warin, K., Limprasert, W., Suebnukarn, S., Jinaporntham, S., Jantana, P., & Vicharueang, S. (2022). AI-based analysis of oral lesions using novel deep convolutional neural networks for early detection of oral cancer. PloS One, 17(8), e0273508. doi:10.1371/journal.pone.0273508

[7] Mentel, S., Gallo, K., Wagendorf, O., Preissner, R., Nahles, S., Heiland, M., & Preissner, S. (2021). Prediction of oral squamous cell carcinoma based on machine learning of breath samples: a prospective controlled study. BMC Oral Health, 21(1), 500. doi:10.1186/s12903-021-01862-z

[8] Alabi, R. O., Bello, I. O., Youssef, O., Elmusrati, M., Mäkitie, A. A., & Almangush, A. (2021). Utilizing deep machine learning for prognostication of oral squamous cell carcinoma-A systematic review. Frontiers in Oral Health, 2, 686863. doi:10.3389/froh.2021.686863

[9] Musulin, J., Štifanić, D., Zulijani, A., Ćabov, T., Dekanić, A., & Car, Z. (2021). An enhanced histopathology analysis: An AI-based system for multiclass grading of oral squamous cell carcinoma and segmenting of epithelial and stromal tissue. Cancers, 13(8), 1784. doi:10.3390/cancers13081784

[10] Jubair, F., Al-Karadsheh, O., Malamos, D., Al Mahdi, S., Saad, Y., & Hassona, Y. (2022). A novel lightweight deep convolutional neural network for early detection of oral cancer. Oral Diseases, 28(4), 1123–1130. https://doi.org/10.1111/odi.13825

[11] Rahman, A.-U., Alqahtani, A., Aldhafferi, N., Nasir, M. U., Khan, M. F., Khan, M. A., & Mosavi, A. (2022). Histopathologic oral cancer prediction using oral squamous cell carcinoma biopsy empowered with transfer learning. Sensors (Basel, Switzerland), 22(10), 3833. https://doi.org/10.3390/s22103833

[12] Torabi, Molook, et al. "Correlation between Clinical and Histopathologic Diagnosis of Oral Potentially Malignant Disorder and Oral Squamous Cell Carcinoma." Pesquisa Brasileira Em Odontopediatria e Clinica Integrada, vol. 21, 2021, p. e0143, doi:10.1590/pboci.2021.068.

[13] Rivera, César, and Bernardo Venegas. "Histological and Molecular Aspects of Oral Squamous Cell Carcinoma (Review)." Oncology Letters, vol. 8, no. 1, 2014, pp. 7–11, doi:10.3892/ol.2014.2103.

[14] Gorbachev, Y., Fedorov, M., Slavutin, I., Tugarev, A., Fatekhov, M., & Tarkan, Y. (2019). Openvino deep learning workbench: Comprehensive analysis and tuning of neural networks inference. In Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (pp. 0-0).

[15] Terada, Tadashi. "Verrucous Carcinoma of the Oral Cavity: A Histopathologic Study of 10 Japanese Cases." Journal of Maxillofacial and Oral Surgery, vol. 10, no. 2, 2011, pp. 148–151, doi:10.1007/s12663-011-0197-x.

[16] Yu, Ryan, et al. "Pseudoglandular Squamous Cell Carcinoma." Cutis; Cutaneous Medicine for the Practitioner, vol. 95, no. 2, 2015, pp. 68, 104–6, https://www.mdedge.com/dermatology/article/97147/dermatopathology/pseudoglandular-squamous-cell-carcinoma.

[17] Raut, T., et al. "Adenoid (Acantholytic) Squamous Cell Carcinoma of Mandibular Gingiva." Case Reports in Dentistry, vol. 2021, 2021, p. 5570092, doi:10.1155/2021/5570092.

[18] Kim, Bo Young, et al. "Sarcomatoid Carcinoma after Radiotherapy for Early-Stage Oral Squamous Cell Carcinoma: Case Report: Case Report." Medicine, vol. 98, no. 27, 2019, p. e16003, doi:10.1097/MD.0000000000016003.

[19] Eguchi, Takanori, et al. “Adenosquamous Carcinoma Development as a Recurrence of Squamous Cell Carcinoma in the Oral Floor: A Case Report: A Case Report.” Medicine, vol. 98, no. 43, 2019, p. e17688, doi:10.1097/MD.0000000000017688.
[20] Gupta, Bhavana, et al. “Basaloid Squamous Cell Carcinoma - A Rare and Aggressive Variant of Squamous Cell Carcinoma: A Case Report and Review of Literature.” National Journal of Maxillofacial Surgery, vol. 9, no. 1, 2018, pp. 64–68, doi:10.4103/njms.NJMS_14_17.
[21] Wang, Yong, et al. “Papillary Squamous Cell Carcinoma Successfully Treated with Bronchoscopic Intratumoral Injections of Cisplatin and Endostar: A Case Report.” The Journal of International Medical Research, vol. 49, no. 9, 2021, p. 3000605211047077, doi:10.1177/03000605211047077.
[22] Alabi, Rasheed Omobolaji, et al. “Machine Learning in Oral Squamous Cell Carcinoma: Current Status, Clinical Concerns and Prospects for Future-A Systematic Review.” Artificial Intelligence in Medicine, vol. 115, no. 102060, 2021, p. 102060, doi:10.1016/j.artmed.2021.102060.
[23] Mentel, Sophia, et al. “Prediction of Oral Squamous Cell Carcinoma Based on Machine Learning of Breath Samples: A Prospective Controlled Study.” BMC Oral Health, vol. 21, no. 1, 2021, p. 500, doi:10.1186/s12903-021-01862-z.
[24] Tseng, Yi-Ju, et al. “Development and Validation of Machine Learning-Based Risk Prediction Models of Oral Squamous Cell Carcinoma Using Salivary Autoantibody Biomarkers.” BMC Oral Health, vol. 22, no. 1, 2022, p. 534, doi:10.1186/s12903-022-02607-2.
[25] Chiesa-Estomba, Carlos M., et al. “Machine Learning Algorithms as a Computer-Assisted Decision Tool for Oral Cancer Prognosis and Management Decisions: A Systematic Review.” ORL; Journal for Oto-Rhino-Laryngology and Its Related Specialties, vol. 84, no. 4, 2022, pp. 278–288, doi:10.1159/000520672.
[26] Alabi, Rasheed Omobolaji, et al. “Deep Machine Learning for Oral Cancer: From Precise Diagnosis to Precision Medicine.” Frontiers in Oral Health, vol. 2, 2021, p. 794248, doi:10.3389/froh.2021.794248.
[27] Musulin, Jelena, et al. “Automated Grading of Oral Squamous Cell Carcinoma into Multiple Classes Using Deep Learning Methods.” 2021 IEEE 21st International Conference on Bioinformatics and Bioengineering (BIBE), IEEE, 2021, pp. 1–6.
[28] Alanazi, Adwan A., et al. “Intelligent Deep Learning Enabled Oral Squamous Cell Carcinoma Detection and Classification Using Biomedical Images.” Computational Intelligence and Neuroscience, vol. 2022, 2022, p. 7643967, doi:10.1155/2022/7643967.
[29] Deif, Mohanad A., et al. “Diagnosis of Oral Squamous Cell Carcinoma Using Deep Neural Networks and Binary Particle Swarm Optimization on Histopathological Images: An AIoMT Approach.” Computational Intelligence and Neuroscience, vol. 2022, 2022, p. 6364102, doi:10.1155/2022/6364102.