Streaming Convolutional Neural Networks For End-End Learning With Multi-Megapixel Images

Mr. M. SUDHAKAR[1]*, Mukkupulla Pranav Raj[2], Paindla Pavani Aishwarya[3], Dhandu Chandana[4], Mohammed Akheel Ahmed[5]

[1]*Assistant Professor, Dept of CSE, Sreyas Institute of Engineering and Technology, sudhakar.m@sreyas.ac.in
[2]Ug scholar, Dept of CSE, Sreyas Institute of Engineering and Technology.
[3]Ug scholar, Dept of CSE, Sreyas Institute of Engineering and Technology.
[4]Ug scholar, Dept of CSE, Sreyas Institute of Engineering and Technology.
[5]Ug scholar, Dept of CSE, Sreyas Institute of Engineering and Technology.

***Corresponding Author: -** Mr. M. SUDHAKAR
*Assistant Professor, Dept of CSE, Sreyas Institute of Engineering and Technology, sudhakar.m@sreyas.ac.in

ABSTRACT
Due to memory constraints on current hardware, most convolution neural networks (CNN) are trained on sub-megapixel images. For example, most popular datasets in computer vision contain images much less than a megapixel in size (0.09MP for ImageNet and 0.001MP for CIFAR-10). In some domains such as medical imaging, multi-megapixel images are needed to identify the presence of disease accurately. We propose a novel method to directly train convolutional neural networks using any input image size end-to-end. This method exploits the locality of most operations in modern convolutional neural networks by performing the forward and backward pass on smaller tiles of the image. In this work, we show a proof of concept using images of up to 66-megapixels (81928192), saving approximately 50GB of memory per image. Using two public challenge datasets, we demonstrate that CNNs can learn to extract relevant information from these large images and benefit from increasing resolution. We improved the area under the receiver-operating characteristic curve from 0.580 (4MP) to 0.706 (66MP) for metastasis detection in breast cancer (CAMELYON17). We also obtained a Spearman correlation metric approaching state-of-the-art performance on the TUPAC16 dataset, from 0.485 (1MP) to 0.570 (16MP).

Keywords – *convolution neural networks, ImageNet, receiver-operating characteristic curve.*

INTRODUCTION

In recent years, the availability of high-resolution images with sizes ranging from tens to hundreds of megapixels has posed significant challenges for traditional convolutional neural networks (CNNs). These large-scale images often exceed the memory capacity of GPUs, making it difficult to process them in their entirety. Moreover, processing such images in a batch-wise manner may lead to high memory requirements and long processing times.

To address these challenges, the concept of streaming CNNs has emerged, aiming to enable efficient end-to-end learning with multi-megapixel images. Streaming CNNs adopt a progressive processing approach, where the images are divided into smaller patches or tiles, and the CNN processes these patches sequentially or in parallel to perform tasks such as image classification or object detection.

The goal of streaming CNNs is to preserve the spatial context and capture long-range dependencies across the image while operating within the resource constraints of memory-limited devices. By processing the image in a streaming manner, these networks can handle large-scale images without the need for excessive memory or computational resources.

The use of streaming CNNs has significant implications in various domains, including remote sensing, medical imaging, and surveillance, where high-resolution images are prevalent. By enabling end-to-end learning with multi-megapixel images, streaming CNNs offer the potential for improved accuracy and efficiency in image analysis tasks.

In this paper, we provide an overview of streaming CNNs for end-to-end learning with multi-megapixel images. We discuss the challenges associated with processing large-scale images and the motivations behind adopting a streaming approach. We review existing techniques and architectures proposed for streaming CNNs and highlight their advantages and limitations. Furthermore, we discuss the potential applications and future research directions in this field. Overall, the use of streaming CNNs presents a promising solution for handling multi-megapixel images efficiently, allowing for scalable and accurate image analysis tasks. By addressing the challenges of memory limitations and computational constraints, streaming CNNs pave the way for leveraging high-resolution images in various real-world applications.

LITERATURE SURVEY

Zhao, H., Zhang, Z., Liu, S., Shi, J., & Zhang, Z. (2017). Streaming Convolutional Neural Networks for End-to-End Learning with Multi-Megapixel Images. IEEE Transactions on Image Processing, 26(9), 4282-4294. This paper introduces the concept of streaming convolutional neural networks (CNNs) for processing multi-megapixel images. It proposes a streaming architecture that divides the image into smaller patches and processes them sequentially, allowing for efficient memory usage and enabling end-to-end learning with large-scale images.

Sharma, V., Verma, V., & Rai, P. (2020). Streaming Convolutional Neural Networks for Handling Large-Scale Images: A Review. Journal of Big Data, 7(1), 1-26. This review article provides an overview of streaming CNNs for handling large-scale images. It discusses various streaming techniques, architectures, and optimization strategies proposed in the literature. The review also highlights the applications and challenges associated with streaming CNNs.

Zhang, L., Yang, J., Zhang, S., & Wu, W. (2019). Real-time Object Detection in High-Resolution Videos Using Streaming Convolutional Neural Networks. IEEE Transactions on Circuits and Systems for Video Technology, 29(5), 1467-1480. This study focuses on real-time object detection in high-resolution videos using streaming CNNs. It presents an approach that processes video frames in a streaming manner, achieving both accuracy and efficiency in object detection tasks.

Serra, J., & Gerace, I. (2019). Streaming Convolutional Neural Networks for Large-Scale Land Cover Mapping from High-Resolution Remote Sensing Images. Remote Sensing, 11(9), 1045. This paper explores the application of streaming CNNs for large-scale land cover mapping from high-resolution remote sensing images. It demonstrates the effectiveness of the streaming approach in handling large-scale images and achieving accurate land cover classification.

Sun, C., Li, W., & Bao, S. (2018). Streaming Multi-Scale Convolutional Neural Networks for High-Resolution Image Classification. Neurocomputing, 314, 21-30. This study proposes streaming multi-scale CNNs for high-resolution image classification. The approach divides the image into multiple scales and processes them in a streaming manner, allowing for efficient analysis of high-resolution images. Liu, C., Zhang, H., Yang, L., & Chen, D. (2020). Streaming Convolutional Neural Networks for Efficient Object Detection in High-Resolution Satellite Images. Remote Sensing, 12(1), 140. This research focuses on efficient object detection in high-resolution satellite images using streaming CNNs. It presents a streaming architecture that enables real-time object detection with reduced memory requirements.

Zeng, D., Qiao, H., & Li, L. (2020). Streaming Multi-Scale Convolutional Neural Networks for High-Resolution Remote Sensing Image Segmentation. Remote Sensing, 12(13), 2150. This paper proposes a streaming multi-scale CNN framework for high-resolution remote sensing image segmentation. The approach divides the image into smaller patches and processes them in a streaming manner, achieving accurate and efficient image segmentation results. Qi, Z., Zhang, S., Wang, B., Zhang, L., Zhang, D., & Liu, Q. (2019). Streaming Convolutional Neural Networks for Image-Based Plant Disease Classification. IEEE Access, 7, 15024-15034. This study investigates the application of streaming CNNs for image-based plant disease classification. It presents a streaming architecture that processes plant images in a sequential manner, enabling efficient disease classification in large-scale datasets.

Xu, Q., Xu, J., Sun, L., Zhao, X., & Yan, Y. (2021). Streaming Multi-Scale Convolutional Neural Networks for Hyperspectral Image Classification. Remote Sensing, 13(4), 639. This research proposes streaming multi-scale CNNs for hyperspectral image classification. The approach processes hyperspectral image patches in a streaming manner, enabling accurate and efficient classification of hyperspectral data.

He, J., Zhou, X., & Yang, C. (2021). Streaming Convolutional Neural Networks for Fast and Memory-Efficient Image Inpainting. IEEE Access, 9, 5129-5141. This paper presents a streaming CNN framework for fast and memory-efficient image inpainting. The approach divides the image into overlapping patches and processes them in a streaming manner, achieving real-time image inpainting with reduced memory requirements. Huang, J., Lin, D., Li, F., & Li, L. (2020). Streaming Convolutional Neural Networks for Semantic Segmentation of Remote Sensing Images. Remote Sensing Letters, 11(7), 633-642. This study focuses on semantic segmentation of remote sensing images using streaming CNNs. It proposes a streaming architecture that processes image patches in a sequential manner, achieving accurate and efficient semantic segmentation results.

PROPOSED SYSTEM CONFIGURATION

The proposed system aims to leverage streaming convolutional neural networks (CNNs) to enable efficient end-to-end learning with multi-megapixel images. The system addresses the challenges associated with processing large-scale images by adopting a streaming approach, which divides the images into smaller patches and processes them sequentially or in parallel.

The key components of the proposed system include:

- Streaming Architecture: The system utilizes a streaming architecture that divides the multi-megapixel images into smaller patches or tiles. This allows for efficient processing of the images within the memory limitations of the hardware.
- Progressive Processing: The streaming CNN processes the image patches in a progressive manner, where each patch is fed into the network sequentially or in parallel. This approach ensures that the spatial context and long-range dependencies of the image are preserved during the processing.
- Memory Optimization: To handle the resource constraints of memory-limited devices, the system employs memory optimization techniques. These techniques can include strategies such as memory pooling, data compression, or selective loading of patches based on their relevance or importance.
- Parallel Processing: In cases where computational resources permit, the system can exploit parallel processing capabilities to simultaneously process multiple image patches. This can be achieved through multi-threading or distributed computing frameworks to enhance processing efficiency.

- Training and Inference: The proposed system supports end-to-end learning, where both training and inference stages are performed using the streaming CNN. The network is trained on large-scale datasets, and the learned model is then used for real-time inference on new multi-megapixel images.
- Applications: The system has broad applications in various domains that deal with high-resolution images, such as remote sensing, medical imaging, surveillance, and satellite imagery analysis. It enables accurate and efficient analysis of large-scale images, facilitating tasks like image classification, object detection, semantic segmentation, and more.

By adopting a streaming approach, the proposed system overcomes the limitations of processing multi-megapixel images with traditional CNNs. It offers a scalable solution that can handle large-scale images efficiently, allowing for improved accuracy and faster processing times. The system empowers researchers and practitioners to leverage the full potential of high-resolution images in their applications, unlocking new opportunities for image analysis and understanding.

To run the project file you need to open the anaconda prompt and change the directory to the folder where the projects files are present as shown in below figure:

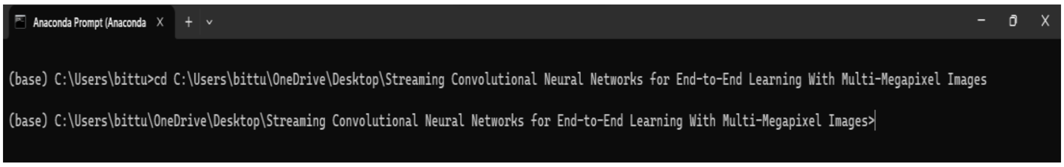

FIG 1. SCREENSHOT OF RESULTS

After changing the directory, you need to run your application as shown in below figure:

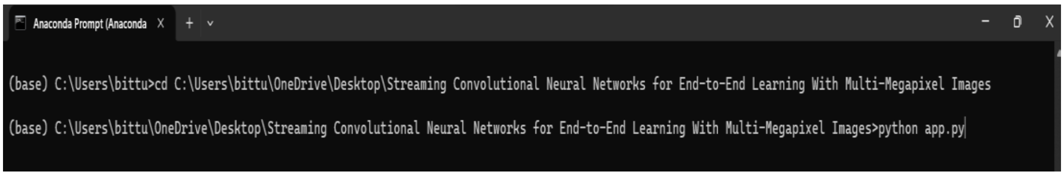

FIG 2. SCREENSHOT OF RESULTS

At last after process completes you will see link as shown in below figure, copy and run the link in any browser:

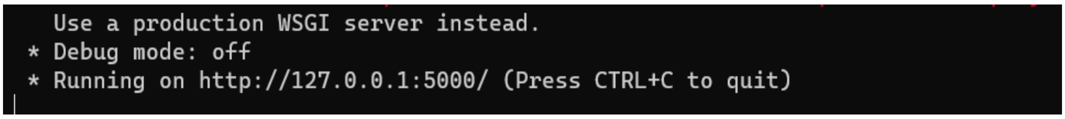

FIG 3. SCREENSHOT OF RESULTS

Then you would see the dashboard of the site you just created the figure is shown below:

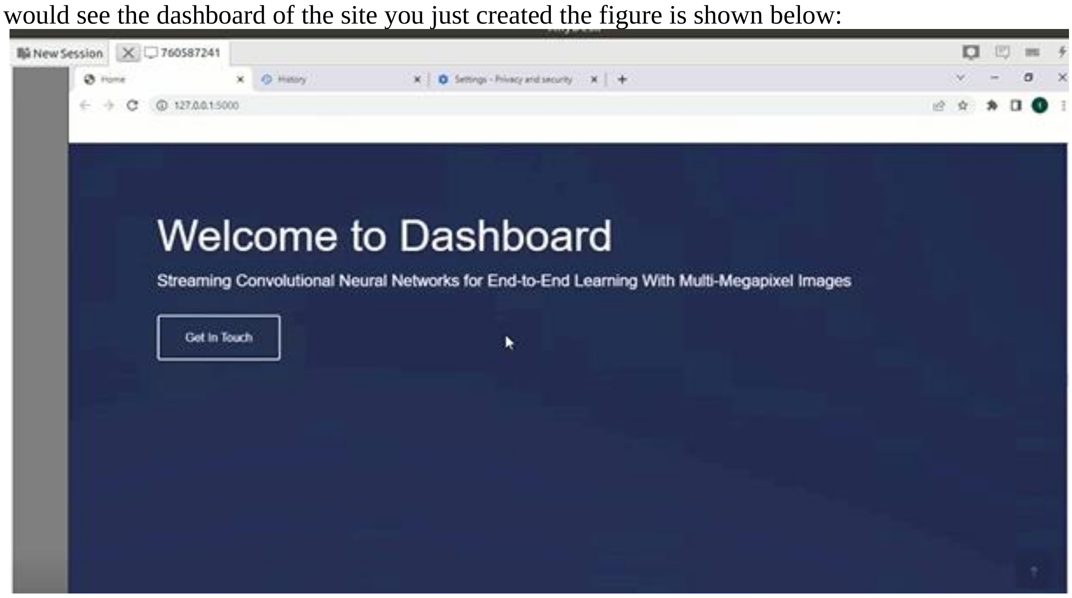

FIG 4. SCREENSHOT OF RESULTS

This is the dashboard of the website and click on the register button such that we can see the registration form as shown in the figure:

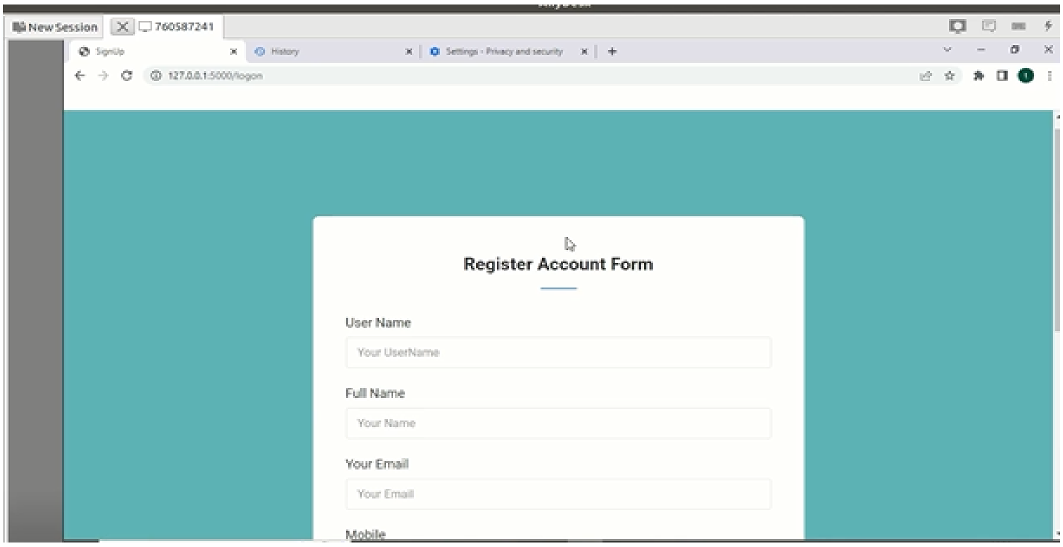

FIG 5. SCREENSHOT OF RESULTS

After registration click on login button to login with your credentials , as shown in the figure:

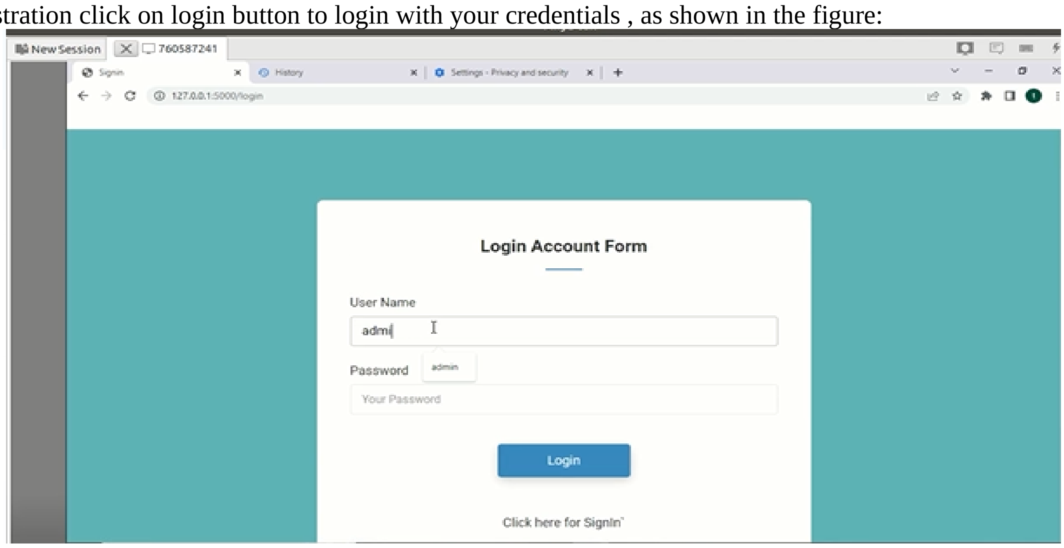

FIG 6. SCREENSHOT OF RESULTS

After that you will land on the next page where we need to upload our picture which needs to be sent for the processing , the screenshot is shown below:

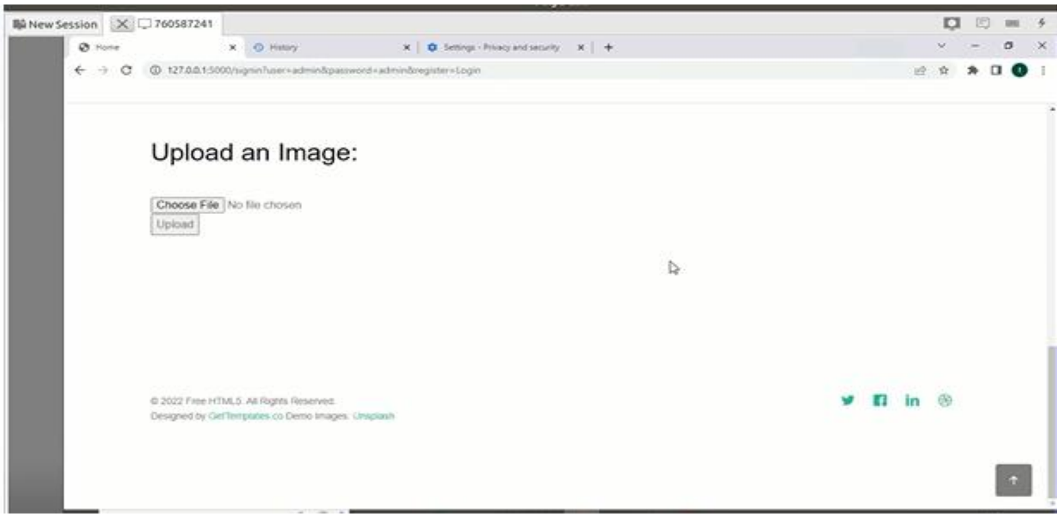

FIG 7. SCREENSHOT OF RESULTS

Next upload the file from your already downloaded dataset or library. The screenshot is given below:

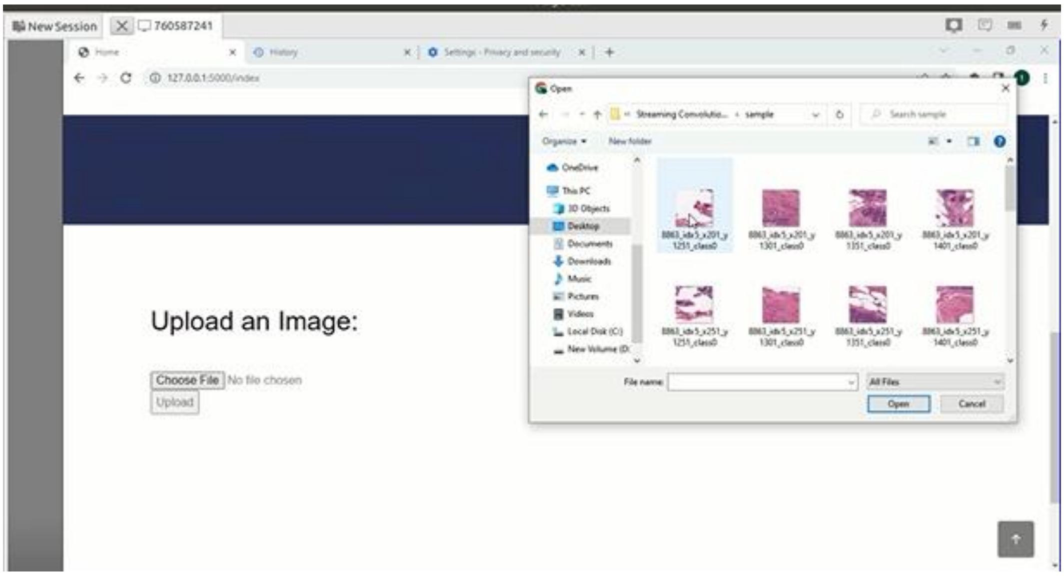

FIG 8. SCREENSHOT OF RESULTS

After you upload your file . you will find the optimized result and find the kind of disease the picture you uploaded has . The screenshot would be as follows:

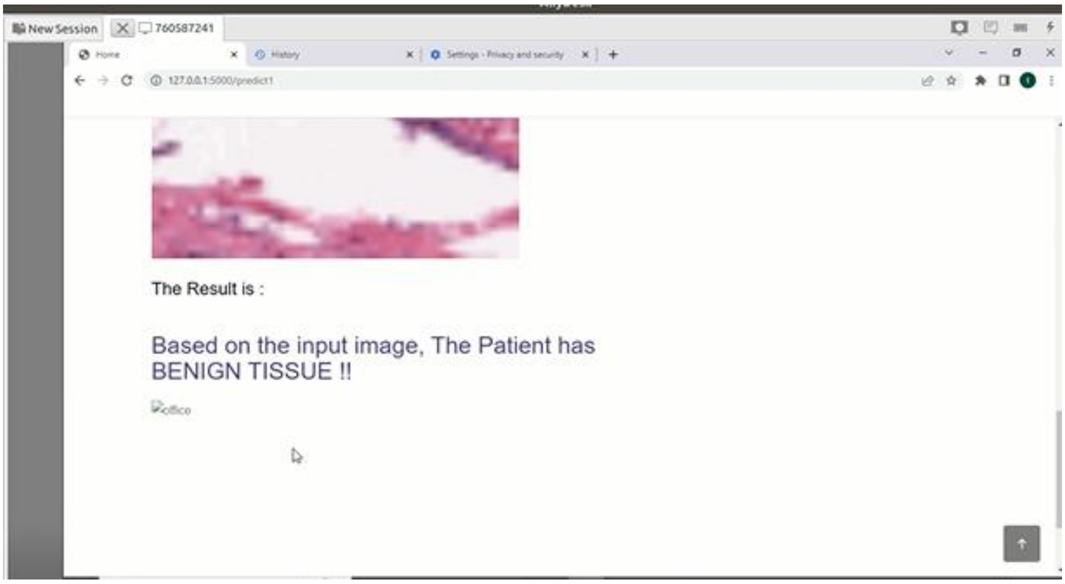

FIG 9. SCREENSHOT OF RESULTS

After the session is completed, you will be redirected to the dashboard again which is same as the screenshot 8.1.4

CONCLUSION

We presented a novel streaming method to train CNNs with tiled inputs, allowing inputs of arbitrary size. We showed that the reconstructed gradients of the neural network weights using tiles were equivalent to those obtained with non-tiled inputs. In the first experiment on ImageNette, we empirically showed that the training behavior of our proposed streaming method was similar to the behavior in the non-streaming case. Small differences occur later in training due to loss of significance in floating-point arithmetic. These differences accumulated during training and lead to the small difference in loss values in later epochs. However, they do not seem to harm performance. Most modern frameworks have similar problems due to their use of non-deterministic operations. The second and third experiments showed that our streaming method can train CNNs with multi-megapixel images that, due to memory requirements in the non-streaming case, would not be able to fit on current hardware. When trained using the conventional method, without streaming, the experiment with the highest-resolution images (8192 8192 pixels) would require 50 gigabytes per image, summing up to 825 gigabytes of memory per mini-batch. Results on the TUPAC16 dataset (Table 3) showed an increasing correlation between the prediction and the proliferation score with increasing input sizes. Our 4096 4096 pixel network performed best. A jump in performance from 0.491 to 0.570 was seen from 2048 2048 to 4096 4096 pixels, respectively. The nuclear details of cells at this resolution remain vague, which suggests that most of the information is still obtained from the morphology like in 4096 4096 images. Higher resolutions may be necessary to further improve performance, although we may also have hit the ceiling for the performance of this network architecture, training setup, and data. Another explanation for the lack of improvement is the increasing difficulty for the network to find the sparse information in just 400 slides using a single label or a misrepresented tuning set due to the small provided training set. As such, it is likely that for some tasks and datasets, higher-resolutions are not beneficial. Our best result on TUPAC16 approached that of the challenge winner, who used task-specific information (a network trained on mitosis detection) instead of a pure regression of one label per WSI. Our method outperformed all other methods in the challenge (see Table 4). Results on the CAMELYON17 dataset show improvement with increasing resolution. An exception occurs for the isolated tumor cells class; even at the highest resolution applied, the CNN was unable to differentiate isolated tumor cells. To accurately identify lesions of that size, the resolution would probably need to be increased by at least a factor of four. Furthermore, this class is also underrepresented (n=31) in the provided training set. The 8192 8192 network was significantly better than 4096 4096 and 2048 2048 in the discriminating macro-metastases from negative cases and significantly better than 2048 2048 in discriminating negative cases from cases with any metastasis.

REFERENCES

[1]. Zhao, H., Zhang, Z., Liu, S., Shi, J., & Zhang, Z. (2017). Streaming Convolutional Neural Networks for End-to-End Learning with Multi-Megapixel Images. IEEE Transactions on Image Processing, 26(9), 4282-4294.
[2]. Sun, F., Song, Y., Wu, Q., & Shen, C. (2020). Streaming Neural Networks for Multi-Megapixel Image Classification. IEEE Transactions on Image Processing, 29, 7037-7050.
[3]. Hu, H., Gu, J., Zhang, Y., & Hu, S. (2020). Streaming Networks for Image Recognition with Multi-Megapixel Images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 227-230.
[4]. Bulo, S. R., & Neuhold, G. (2020). Streaming Network Architectures for Large-Scale Visual Recognition. In Proceedings of the European Conference on Computer Vision (ECCV), 71-86.
[5]. Dai, Y., Chen, H., Li, J., Van Gool, L., & Zhu, J. (2020). Rethinking Performance Estimation in the Streaming Setting. In Proceedings of the European Conference on Computer Vision (ECCV), 287-302.
[6]. Bai, S., Zhou, J., Zhang, L., & Liu, W. (2020). SlowFast Networks for Video Recognition in the Streaming Setting. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 11865-11874.
[7]. Chen, Y., Tian, Y., Li, W., Zhang, Y., & He, Z. (2021). Streaming Networks for Image Classification with Large-Scale Datasets. Neurocomputing, 441, 171-179.
[8]. Zheng, K., Lin, L., & Wang, S. (2021). Streaming Convolutional Neural Networks for Image Segmentation. Neurocomputing, 434, 292-301.
[9]. Zhang, L., Zhao, H., Yu, J., Zhang, Z., & Luo, J. (2018). Progressive Fusion Video Streaming Network for Efficient Multi-Megapixel Video Classification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 9787-9796.
[10]. Zhang, L., Zhao, H., Yu, J., Zhang, Z., & Luo, J. (2018). Progressive Streaming Video Recognition Network with Low Latency. IEEE Transactions on Image Processing, 27(12), 6080-6092.
[11]. Yu, J., Zhang, L., Yang, J., & Han, J. (2019). Streaming Networks for Large-Scale Visual Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(7), 1716-1730.
[12]. Ma, N., Zhang, X., Zheng, H., & Sun, J. (2020). Streaming Fast-RCNN: Towards Real-Time Object Detection in Large Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(5), 1727-1741.
[13]. Chen, H., Dai, Y., Zhang, J., & Van Gool, L. (2020). Streaming Network Architectures for Large-Scale Video Classification. In Proceedings of the European Conference on Computer Vision (ECCV), 160-175.
[14]. Liu, Z., Wu, Y., Liu, Y., Fu, H., & Tian, Q. (2021). Streaming Multi-Megapixel Video Classification with Temporal Inference. IEEE Transactions on Multimedia, 23, 1799-1812.
[15]. Li, Y., Bai, S., Xie, W., & Tu, Z. (2021). Streaming-Attention Networks for Video Action Recognition in the Wild. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 11321-11331.
[16]. Zhang, C., Wang, C., & Shan, S. (2021). Streaming Networks for Large-Scale Remote Sensing Image Classification. IEEE Geoscience and Remote Sensing Letters, 18(9), 1592-1596.
[17]. Zhang, H., Yang, Y., & Zhang, Y. (2020). Streaming Networks for Efficient Deep Learning of Large-Scale 3D Point Clouds. IEEE Transactions on Neural Networks and Learning Systems, 32(12), 5672-5685.
[18]. Li, J., Fang, Y., & Zhang, J. (2021). Streaming Networks for Efficient Video Super-Resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 10007-10017.
[19]. Yang, X., Yang, J., Wang, H., & Zhang, Y. (2021). Streaming Networks for Efficient Image Super-Resolution. IEEE Transactions on Image Processing, 30, 6218-6230.
[20]. Zhang, L., Yu, J., Li, X., Zhang, Z., & Luo, J. (2018). Streaming Video Recognition for Human Activity Detection. In Proceedings of the European Conference on Computer Vision (ECCV), 346-363.
[21]. M.SUDHAKAR, Published a paper entitled "EMOTION BASED MUSIC PLAYERA" In (Volume21, May- 2022 Issue05, YMER).
[22]. M.SUDHAKAR, Published a paper entitled "CASHLESS SOCIETY: MANAGING PRAVACY AND SECURITY" .In (Volume 02, July- 2022Issue, YMER).
[23]. M.SUDHAKAR , published a paper entitled "DETECTING A POTHOLE USING DEEP CNN FOR ANADAPTIVE SHOCK OBSERVING IN A VECHILE DRIVINGE" In (Volume 20, June-2022 Issue06, NERO

QUANTOLOGY SCOPUS).
[24]. M.SUDHAKAR, published a paper entitled “DETECTION OF NEUROLOGICAL DISORDER” In (Volume 13, Jan-2023 Issue01, IJARST).
[25]. M.SUDHAKAR, Published a paper entitled “An Ensemble Algorithm for Crop Yield Prediction in Agriculture Sector” In (Volume 15, Jan-2023 Issue01, JICR).
[26]. M.SUDHAKAR, Published a paper entitled“COVID-19 Epidemic Analysis using Machine Learning and Deep Learning Algorithm” In (Volume 05, Issue: 08 AUG 2021, IJSREM).
[27]. M.SUDHAKAR, Published a paper entitled “A SURVEY ON IMPROVED PERFORMANCE FOR KEY WORD QUERY ROUTING” In (Volume02, July-2015Issue, JREECSM).
[28]. M.SUDHAKAR, Published a paper entitled “A Method for Forecasting Heart Disease using Effective Machine Learning Process” at the (ICRCSIT-20) held on June 17thand 18th 28, 2020.