International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0260
e ISSN: 2584-2854
Volume: 03
Issue: 05 May 2025
Page No: 1611 - 1618

Convolutional Neural Network Competence in Image Analytics for Tree Enumeration

Dr Anitha M[1], Swaroop N[2], Vinayak Benal[3], Tarun D N[4], Swapnil Nayak[5]
[1]Associate professor, Dept. of ECE, BMS Institute of Technology & Management., Bangalore, India
[2,3,4]UG Scholar, Dept. of ECE, BMS Institute of Technology & Management., Bangalore, India
Emails: *anitham@bmsit.in[1], swaroopnataraj56@gmail.com[2], vinayakbenal131@gmail.com[3], tarundn0@gmail.com[4], swapnilnayak.1234@gmail.com[5]*

Abstract
Tree identification and counting are crucial tasks for forest management, conservation, and planning. Traditional methods, such as manual surveys, are time-consuming, costly, and prone to errors. This paper aims at developing an efficient system for an automated tree detection and counting with the model of YOLO (You Only Look Once). Proposed approach uses annotated dataset from Rob flow and geographical data to improve accuracy and efficiency in tree detection. This YOLO method is optimized for the precise instance of tree detection. This technique comprises new approaches to candidate positive sample selection, which have a new design specific to small and medium-sized tree. The parameters of the YOLOv8 model will be optimized to have a high F1 score and recall rate so that the correct trees are identified. The evaluation metrics are F1 score, precision, and recall to balance the false positives with false negatives. This will streamline the environmental assessments by automatically identifying and counting trees, thereby giving a clear indication of how well managed the forest is. Such innovation demonstrates how integration of machine learning techniques with advanced data sources could be very important in enhancing the tree detection and counting task.

Keywords: *Image Analytics, Remote Sensing, Convolutional Neural Network, Forest Management, Tree Detection.*

1. Introduction
Sustainable tree resource management is the key to mitigating climate warming, fostering a green economy, and protecting valuable habitats. Detailed knowledge about tree resources is a prerequisite for such management but is conventionally based on plot-scale data, which often neglects trees outside forests .In urban environments, the technology can aid in city planning and green space management by providing precise data on tree canopy coverage and its impact on urban ecosystems. Many improved trees counting model and the model's adaptability for different applications, such as species identification and growth monitoring, extends its utility beyond traditional forestry, supporting broader ecological and environmental research. Accurate tree counting is essential for a responsible forest land diversion in development projects. Traditional manual surveys are very slow, costly, and prone to human errors. The primary objective in this sustainable development goal is to develop robust system that identifies and classifies trees. Advanced computer vision techniques are integrated with machine learning models that are used to analyze the images. The results are impressive, significantly accelerating tree enumeration and eliminating resource intensive manual labor. The solution demonstrates precision with a minimal false negative and positives, and efficiently detects trees, offering a comprehensive view of the forested area. This project's significance lies in its contribution to responsible and sustainable land development practices by considering the tree population. By automating tree enumeration, it

e ISSN: 2584-2854
Volume: 03
Issue: 05 May 2025
Page No: 1611 - 1618

allows stakeholders with a timely, informed decisions making from precise data for land usage, conservation, and environmental impact assessments[1]. The solution strikes a balance between ecological preservation and development, optimizing resource allocation while also minimizing environmental impact in forest region[2]. Detailed knowledge of forests at national and regional scales is commonly obtained from inventories such as the national forest inventories (NFI). Here, variables such as tree diameter, height, species, growth, and mortality are recorded during repeated census on a representative sample of widely distributed plots [5–8]. Inventories provide essential information on forest biomass stocks used for climate treaties and carbon accounting but are time-consuming, labor-intensive, and limited to plot scale, and the methods, and degree to which monitoring of trees outside forests is conducted, vary substantially across countries. The cutting-edge image analytics solution revolutionizes forest land diversion by enabling ecologically conscious decision-making and efficiency. It answers the critical need for accurate tree counting in the face of developmental challenges and fosters environmental stewardship and responsible land usage leading to responsible land development culture. [3][1]. Integration with Geographic Information Systems (GIS) tool for detailed spatial analysis, mapping, and reporting is discussed in [3]. Remote Sensing combine with high-resolution satellite imagery and LiDAR data for enhanced tree enumeration and also in forest structure analysis [4]. Use drones and UAVs for high-resolution aerial surveys to improve real-time monitoring and data collection [5]. Few models include tree species classification for biodiversity monitoring and forest health assessments[7] .Tools are developed for tree growth monitoring and detecting signs of disease or pests to support proactive forest management [8]. Few models are involved for urban and suburban environments to assist in city planning, green space management, and analyzing tree canopy coverage[9,10]. Object recognition using convolutional neural networks (CNNs) have made a significant progress in recent years in forest management [11], but there is still much to be done. You Only Look Once (YOLO) is a state-of-the-art model for detecting objects in videos and images with a very high accuracy and greater precision. Unlike traditional methods that apply detection over multiple regions of an image, YOLO does this in a single pass, which makes it fast, efficient and suitable for the project. This article is structured as follows. Section 2 introduces theory of Convolutional neural networks which are used in automating the process of identifying and counting trees. Tree counting and crown segmentation Methods with YOLO8 which is suitable to both computational and visualization tasks are presented in Section 3. Section 4 includes the simulation results and analyses. Conclusions are given in Section 5.

2. Model of CNN Competence in Tree Detection

Computer vision consists of various fields such as image recognition, object detection in images or video, image processing, video analysis, and more. Object detection is the most depth field of computer vision technology. Object recognition is a complex task of computer vision that involves identification and classification of objects in an video and image. Due to their ability to learn complex features and patterns from raw image and video data, the most effective and popular approach to object recognition has been with convolutional neural networks (CNNs). CNNs have revolutionized the field of computer vision and led to a significant advance in object recognition. CNNs are composed of multiple layers of interconnected neurons that learn and extract features from images or videos. The first layer of a CNN typically consists of a set of filters that convolve over the input image or video, producing feature maps that highlight important patterns such as edges and corners. The subsequent layers in the network learn more complex features by convolving over the feature maps produced by the previous layers. CNN is a deep learning method, recently it has stepped forward and there has been a vivid development in the field of computer vision such as an object detection, image segmentation, recognition, and captioning [12][13]. It is biologically inspired by the human brain. Convolutional neural nets and other related architectures under the deep learning umbrella are best comparable to the neural networks that exist

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0260
e ISSN: 2584-2854
Volume: 03
Issue: 05 May 2025
Page No: 1611 - 1618

in the human brain. Just like the structure and the operations of the neurons in the human brain through the human visual cortex, through the deployment of hierarchical multi-layers networks. The deep neural networks have revealed very effective way to learn various schemes of feature and pattern recognition and representation from the training data. Its operations are automatic and features engineering tasks that can be resolved fast and a reliable way. They are very much capable of finding and effectively using specific peculiarities of image object in case a massive training dataset is provided. In the early 90s, CNN was first used for the purpose of recognizing handwritten digits [14,15]. Later on, in the early 2010s a major development was made by releasing Alex Net [16]. There is a basic principle that needs to be considered for special case for multilayer perceptron where every neuron is connected to the receptive field located in forward-face. Moreover, the neurons belong to each layer in the network share the same weights as the other neurons (Figure 1).

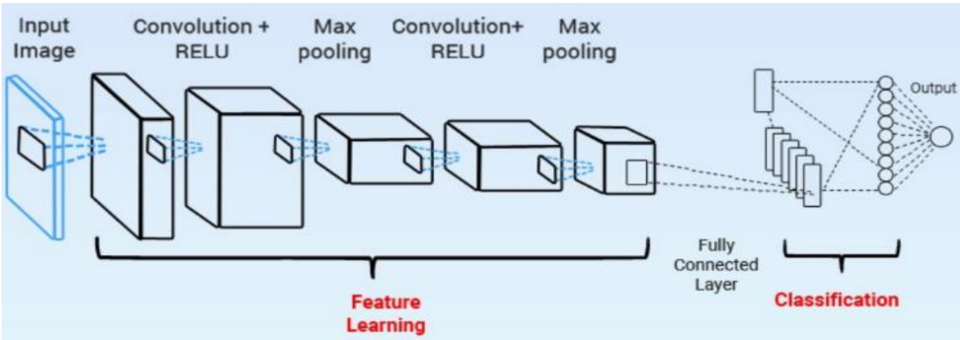

Figure 1 **Convolutional Neural Network**

An example of object recognition using CNNs is image classification. In this work, a CNN is trained on a large dataset of images with associated labels of trees. The network learns to recognize patterns and features in the images that are associated with different object categories with labels. Once trained, the model can classify new images into the appropriate object category with a high degree of accuracy and precision (Figure 2).

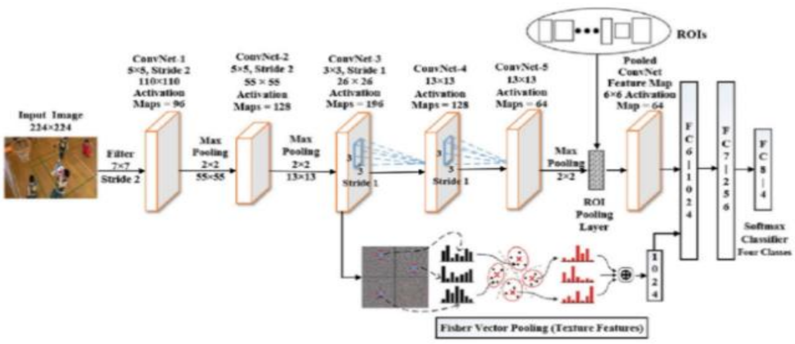

Figure 2 **Image Classification using CNN**

Although object recognition using convolutional neural networks (CNNs) has made a significant progress in recent years, there are still several challenges and limitations associated with this approach. Some of the main problems include the following:

- **Data bias:** CNNs use a lot of training data to learn and classify objects. If, however, the training data is biased to favour one particular kind of object or environment, then CNN might fail in a variety of settings. Errors and misclassifications result because of such

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0260
e ISSN: 2584-2854
Volume: 03
Issue: 05 May 2025
Page No: 1611 - 1618

errors and especially when new, unseen objects appear in the setting where CNN was first introduced.

- **Overfitting:** CNNs easily overfit the training data; the network specializes too much and fails to generalize to new, unseen data. Techniques such as regularization and data augmentation can help with that.
- **Computational complexity:** CNNs are computationally expensive and require more computational power for training and execution. This often makes it challenging to scale CNNs up to large-Scale applications.
- **Interpretability:** CNNs often are black boxes meaning it is usually impossible to understand how a network has arrived at its decision, which can prove problematic, for example medical image analysis where these properties are to be valued as important.

These range from data bias and overfitting to computational complexity and interpretability. The solutions to these problems will be important for further development and growth in object recognition using CNNs.

3. Object Recognition and Classification Models in Image Analytics

3.1. Improving Object Recognition and classification using Convolutional Neural Network

Object recognition and classification models are essential components of image analytics, which enable machines to identify, detect, and categorize objects in images. Object recognition using CNN has become one of the most popular and effective approaches in computer vision and object detection. The general techniques to improve object recognition using neural network are transfer learning, data augmentation, attention mechanisms, network architecture, and optimization algorithms. Integrating these techniques can continue to improve accuracy and efficiency in object recognition using CNN.

3.1.1. Exploring CNN Model Variants

Computer vision has been revolutionized by Convolution Neural Network(CNN) models, which power everything from medical imaging to driverless cars. Notwithstanding their achievements, they still have drawbacks, including the requirement for sizable labeled datasets, sensitivity to changes in the data, high processing requirements, and restricted generalization to various contexts. CNNs also have trouble differentiating fine-grained objects and recognizing cluttered or poor-quality images. Resolving these problems is essential to creating object recognition models that are reliable and effective.

3.1.2. Single-Shot Multibox Detector (SSD)

One well-liked convolutional neural network (CNN) architecture for real-time object detection is the Single-Shot Multibox Detector (SSD). In order to detect objects across many dimensions in a single forward run through the network, it converts bounding box predictions into a collection of boxes with variable sizes and aspect ratios. Applications needing real-time performance can benefit from the SSD's reputation for speed and accuracy. It performs better than models like YOLO for smaller items and is excellent at detecting objects in a single stage. Performance in complicated scenarios may be hampered by SSD's drawbacks, which include its propensity to suffer from subpar feature extraction in shallow levels and its loss of feature richness in deeper layers.

3.2. Deep Learning Methods for Tree counting and crown segmentation Methods Using YOLOv8

Many methods can be tailored depending on the specific needs of the study (e.g., forest density, tree species, image quality). Often, hybrid approaches combining multiple methods (such as CNNs with LiDAR) produce the best results [17]. A convolutional neural network (CNN) is used in the single-stage object recognition framework You Only Look Once (YOLO) to concurrently forecast bounding boxes, calculate class probabilities, and give confidence scores to objects in an image. YOLO, which was created for real-time applications, can quickly detect and recognize objects, including pedestrians, by processing complete photos in a single pass. Because of its effectiveness, it is especially well-suited for applications where speed is crucial, such video surveillance and autonomous driving. YOLO has certain drawbacks despite its

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0260
e ISSN: 2584-2854
Volume: 03
Issue: 05 May 2025
Page No: 1611 - 1618

benefits. Generally speaking, its accuracy is lower than that of models like SSD, and because of its limited capacity to suggest numerous bounding boxes per object, it has trouble handling dense barriers like flocks of birds. Furthermore, YOLO performs poorly when it comes to small object identification. YOLOv8 follows the network architecture of YOLOv5. YOLOv8-seg is a segmentation branch of the YOLOv8 model. Training YOLOv8 on a custom dataset involves preparing a dataset specific to the detection task and configuring the model parameters (Figure 3).

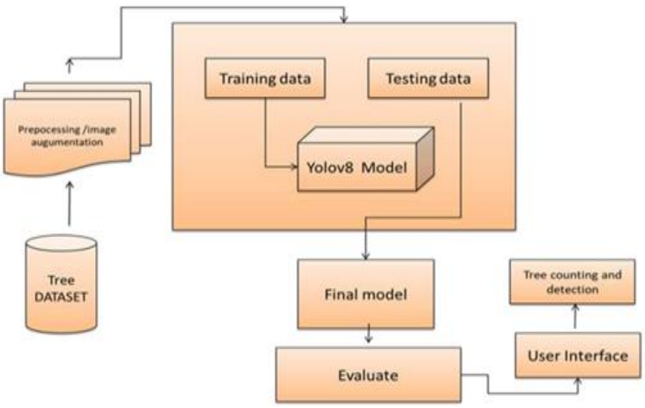

Figure 3 Project Architecture

The process begins with collecting and annotating images that represent the objects of interest, ensuring the model can learn to identify and locate these objects in different conditions.

4. Results, Study Site, Data Collection and Simulation Environment

Throughout project period, the annotation and preprocessing procedure was carried out, and frequent evaluations were made to ensure labelling consistency and correctness. Following annotation, the labelled data was exported in formats that worked with popular machine learning frameworks like Py-Torch and TensorFlow. To verify the accuracy of the annotations, a portion of the data underwent additional validation. Even though Roboflow greatly expedited the process, human labelling presented difficulties, especially in photos with dense tree clusters. To overcome these obstacles, Roboflow's automated features—such as augmentation tools and annotation templates—were used to cut down on the time and effort needed for labelling. Notwithstanding these challenges, the finished dataset turned out to be a useful tool for achieving the project's goals and allowed precise tree counting using machine learning methods. Performance Evaluation of YOLO Models for Tree Enumeration. The segmentation and performance of the YOLOv8 model configurations were evaluated using masked mean average precision (mAP), recall (R), and F1-score. Precision evaluates the accuracy of the predicted positive detections, calculated as

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}}$$

TP counts the correctly identified trees, while FP indicates non-tree objects that are incorrectly identified as trees. The missing trees that the model failed to detect are indicated by FN. The detection model's performance across different threshold levels is measured by AP, which is the region enclosed by the recall rate, precision, and horizontal axis. On the other hand, the model's total performance in tree detection is summarized by mAP, a single, integrated performance parameter. A thorough picture of the model's performance in identifying and counting trees in various settings may be obtained by averaging the AP across all classes. A performance parameter called recall shows how can recognize. It is calculated as

$$\text{Recall} = \frac{\text{True positives (TP)}}{\text{True positives (TP)} + \text{False negatives (FN)}}$$

Recall, also known as sensitivity or true positive rate, measures a model's ability to correctly identify all actual positive cases. It is calculated as the ratio of true positives to the total number of actual positives (true positives + false negatives). A high recall means the model successfully minimizes false negatives, which is critical in scenarios like disease diagnosis or fraud detection, where missing a positive case could have significant consequences. The F1-score measures both the precision and recall of the model

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0260
e ISSN: 2584-2854
Volume: 03
Issue: 05 May 2025
Page No: 1611 - 1618

to compute a single score that represents the model's performance and is calculated as:

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

The F1 score is like a referee balancing two players—precision and recall—to ensure fair play. It gives you a single number that tells how well your model is doing overall when both avoiding false alarms (precision) and catching all the true positives (recall) are important. If your model is great at one but terrible at the other, the F1 score will call it out. Think of it as a handshake between precision and recall, ensuring they work well together rather than focusing on just one (Figure 4). These are the specific metrics used to evaluate the model. In object detection, common metrics include:

Precision: The proportion of true positive detections out of all positive detections made by the model.

Recall: The proportion of true positive detections out of all actual positives in the data.

mAP50: Mean Average Precision at IoU threshold 0.5.

mAP50-95: Mean Average Precision averaged over IoU thresholds from 0.5 to 0.95.

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

Figure 4 Model Training Result

Table 1 Model Training Evaluation Metric

Precision	0.854
Recall	0.808
mAP50	0.863
mAP50-95	0.402
Fitness	0.414

Conclusion

The application of YOLOv8 for tree enumeration represents a significant advancement in the field of forest management and ecological monitoring (Table 1). Through the use of advanced object detection techniques, the project has demonstrated the potential of machine learning models to accurately detect and count trees in high-resolution aerial imagery. The YOLOv8 model's strong performance in precision and recall underscores its effectiveness in various forest conditions, showcasing its capability to handle both dense and sparse tree populations. The practical implications of this technology are profound. In forest management, accurate tree enumeration enables better planning and resource allocation, allowing for more informed decisions regarding conservation and sustainable forestry practices. The model's ability to process large volumes of imagery in real-time supports efficient monitoring and management of vast forested areas. Furthermore, the integration of YOLOv8 with Geographic Information Systems (GIS) and remote sensing technologies opens up new avenues for detailed spatial analysis, enhancing our understanding of forest health and structure. By advancing these capabilities, the project not only improves forest management practices but also contributes to the preservation and enhancement of natural environments in various real-world contexts. Future work is to Create user-friendly interfaces and educational resources for non-experts to interact with and understand the model's outputs.

Reference

[1]. M. Roy and J. Bhaduri, "DenseSPH-YOLOv5: An automated damage detection model based on DenseNet and Swin-Transformer prediction head-enabled YOLOv5 with attention mechanism," Adv. Eng. Inf., vol. 56, p. 102007, 2023, doi: 10.1016/j.aei.2023.102007.

[2]. S. Jamil and A. M. Roy, "An efficient and robust phonocardiography (pcg)-based valvular heart diseases (vhd) detection framework using vision transformer (vit)," Comput. Biol. Med., vol. 158, p. 106734, 2023, doi: 10.1016/j.compbiomed.2023.10673.

[3]. Nofirman, & Ahmada, Naufal & Fauzan, Tribowo. (2024). Integration of Geographic Information Systems and Spatial Data Analysis in Location Decision Making for Manufacturing Industries. International Journal Software Engineering and Computer Science (IJSECS). 4. 196-209. 10.35870/ijsecs.v4i1.2027.

[4]. George-Chacon, S.P.; Dupuy, J.M.; Peduzzi, A.; Hernandez-Stefanoni, J.L. Combining High Resolution Satellite Imagery and Lidar Data to Model Woody Species Diversity of Tropical Dry Forests. Ecol. Indic. 2019, 101, 975–984.

[5]. Han, Pengcheng, Cunbao Ma, Jian Chen, Lin Chen, Shuhui Bu, Shibiao Xu, Yong Zhao, Chenhua Zhang, and Tatsuya Hagino. 2022. "Fast Tree Detection and Counting on UAVs for Sequential Aerial Images with Generating Orthophoto Mosaicing" Remote Sensing 14, no. 16: 4113. https://doi.org/10.3390/rs14164113.

[6]. Abdollahnejad, Azadeh, and Dimitrios Panagiotidis. 2020. "Tree Species Classification and Health Status Assessment for a Mixed Broadleaf-Conifer Forest with UAS Multispectral Imaging" Remote Sensing 12, no. 22: 3722. https://doi.org/10.3390/rs12223722.

[7]. Vanguri et al., 2024 R. Vanguri, G. Laneve, A. Hościło Mapping forest tree species and its biodiversity using EnMAP hyperspectral data along with Sentinel-2 temporal data: an approach of tree species classification and diversity indices Indic., 167 (2024), 10.1016/J.ECOLIND.2024.112671.

[8]. Deepak Kumar Mahanta, Tanmaya Kumar Bhoi, J Komal, Ipsita Samal, Andrea Mastinu, Spatial, spectral and temporal insights: harnessing high-resolution satellite remote sensing and artificial intelligence for early monitoring of wood boring pests in forests, Plant Stress, Volume 11, 2024, 100381, ISSN 2667-064X, https: //doi.org/ 10.1016/j. stress. 2024.100381.

[9]. Luisa Velasquez-Camacho, Maddi Etxegarai, Sergio de-Miguel, Implementing Deep Learning algorithms for urban tree detection and geolocation with high-resolution aerial, satellite, and ground-level images. Volume 105, 2023, 102025, ISSN 0198-9715, https://doi.org/10.1016/j.compenvurbsys.2023.102025.

[10]. Shivesh Kishore Karan, Bjørn Tobias Borchsenius, Misganu Debella-Gilo, Mapping urban green structures using object-based analysis of satellite imagery. January 2025 DOI: 10.1016/j.ecolind.2024.113027

[11]. Kattenborn, T., Leitloff, J., Schiefer, F., Hinz, S., 2021. Review on convolutional neural networks (CNN) in vegetation remote sensing. ISPRS J. Photogramm. Remote Sens. 173, 24–49. https: //doi.org /10.1016/ j.isprsjprs.2020.12.010.

[12]. Chauhan, R., Ghanshala, K. K. & Joshi, R. C. Convolutional neural network (CNN) for image detection and recognition. in 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC) 278–282 (IEEE, 2018). https://doi.org/10.1109/ICSCCC.2018.8703316.

[13]. Schmidhuber, J., Deep learning in neural networks: An overview. Neural Networks, 2015. 61: p. 85-117.

[14]. M. Jain, G. Kaur, M. P. Quamar, and H. Gupta, "Handwritten digit recognition using CNN," Proc. Int. Conf. Innov. Pract. Technol. Manag. ICIPTM 2021, pp. 211–215, Feb. 2021, doi: 10.1109 /ICIPTM52218.2021.9388351.

[15]. LeCun, Y., et al. Handwritten digit recognition with a backpropagation network. in Advances in neural information processing systems. 1990

[16]. Krizhevsky, A., I. Sutskever, and G.E.

Hinton. Imagenet classification with deep convolutional neural networks. in Advances in neural information processing systems. 2012.

[17]. Janne Mäyrä, Sarita Keski-Saari, Sonja Kivinen Tree species classification from airborne hyperspectral and LiDAR data using 3D convolutional neural networks,Remote Sensing of Environment, Volume 256, 2021, 112322, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2021.112322.

[18]. NP Jouppi, C Young, N Patil, D Patterson, G Agrawal, R Bajwa, S Bates, S Bhatia, N Boden, A Borchers, R. Boyle, In-datacenter performance analysis of a tensor processing unit, in: Proc. of the 44th Annual International Symposium on Computer Architecture, 2017, pp. 1–12, https://doi.org/10.1145/3140659.3080246.