

Quantized Mushroom Classification

Nina Immenroth¹, Luke Richard¹¹Berliner Hochschule für Technik

Problem

- Mushroom foraging requires accurate species identification, but many species have highly similar visual appearances
- Misidentification poses safety risks, particularly when distinguishing edible from poisonous mushrooms
- We aim to identify 50 mushroom species from images and provide top-5 predictions with confidence estimates and edibility information
- Offline usability is essential in forest environments with limited or no cell service
- Model compression via weight quantization is explored to enable efficient on-device deployment



Related Work

Fine-grained image classification (He et al., 2016)

- widely studied using deep CNNs
- models such as ResNet and EfficientNet show strong performance on visually similar categories through transfer learning and data augmentation

EfficientNet & EfficientNet-V2 (Tan & Le, 2019; Tan & Le, 2021)

- achieve state-of-the-art accuracy-efficiency trade-offs by jointly scaling network depth, width, and resolution
- well-suited for deployment under computational constraints

Automated mushroom identification

- prior work has explored mushroom classification using fine-tuned CNNs (Jacob, M.S. et al., 2025)
- recent research into model-compression by Adhikari et al. in 2024 → 7 edible species

Model compression & quantization (Jacob, B. et al., 2018; Rokh et al., 2023)

- weight quantization is an established technique for reducing model size and inference latency
- often minimal accuracy degradation, enabling deployment on resource-constrained devices

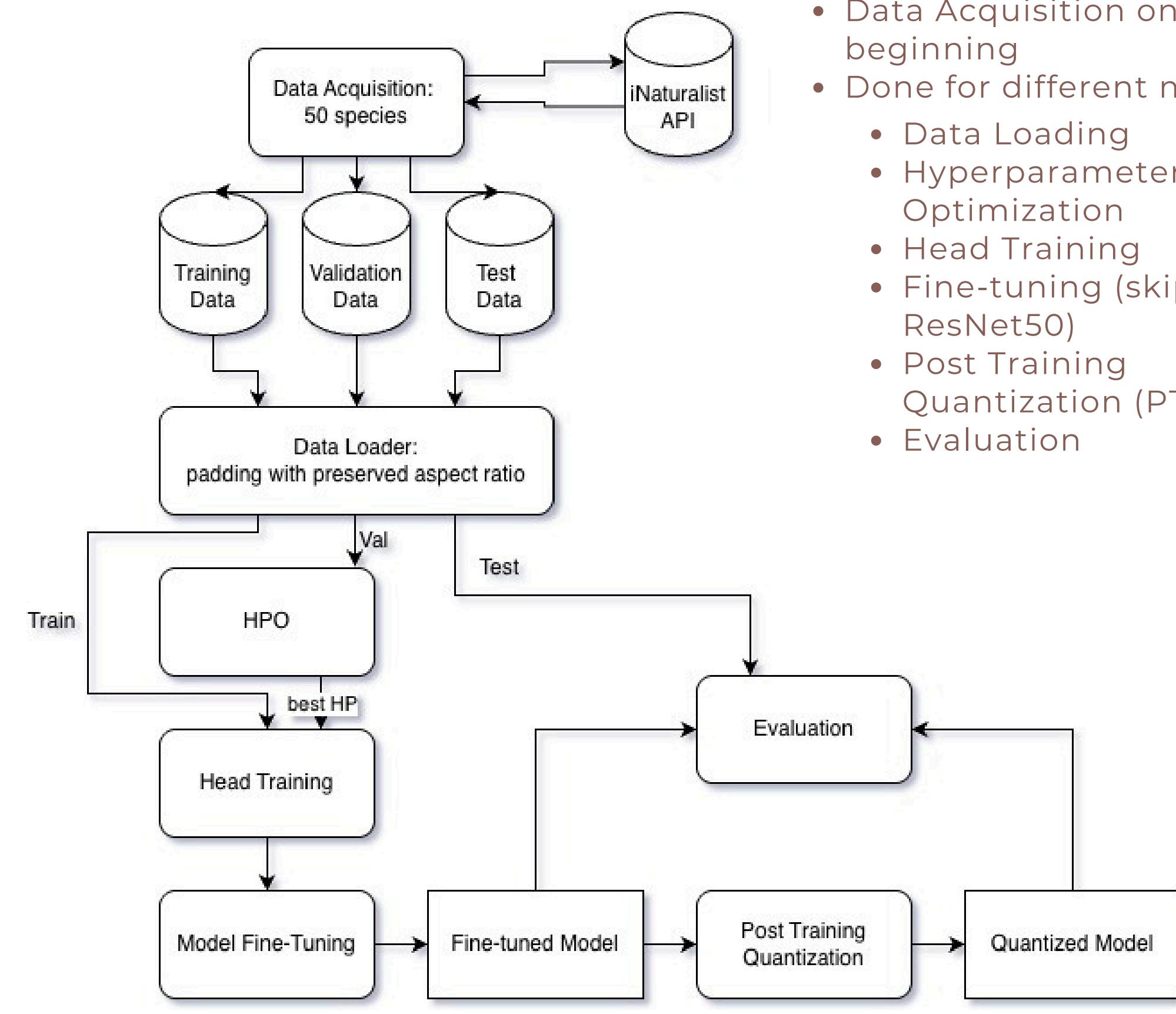
Our Approach

- Data: 50 mushroom species collected from iNaturalist via API; images in low resolution to reflect realistic deployment constraints
- Preprocessing: Aspect-ratio-preserving resizing with padding, followed by data augmentation (random flips, rotations, zoom, and contrast) to improve robustness
- Models: Fine-tuned EfficientNet-V2-M as the primary model, with ResNet as a baseline
- Training: Transfer learning with staged fine-tuning and hyperparameter optimization
- Compression: Post Training Quantization (PTQ) applied to EfficientNet-V2-M to reduce model size and enable fast CPU inference
- Metrics: Top-5 species predictions, edible vs. poisonous classification, and confidence scores to handle ambiguity
- Contribution: Systematic evaluation of accuracy, uncertainty, and efficiency trade-offs under model compression, improving upon prior cloud-dependent mushroom identification systems

References

- Adhikari, B., Hossain, M. M., Akter, S., & Ahmed, M. S. (2024, June). MushroomDetect: An Investigation into Local Mushroom Classification in Bangladesh Using Transfer Learning. In 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT) (pp. 1-6). IEEE.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
- Jacob, B., Kluyz, S., Chen, B., Zhu, M., Tang, M., Howard, A., ... & Kalenichenko, D. (2018). Quantization and training of neural networks for efficient integer-arithmetic-only inference. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2704-2713).
- Jacob, M. S., Xu, A., Qian, K., Qi, Z., Li, X., & Zhang, B. (2025). Artificial Intelligence in Edible Mushroom Cultivation, Breeding, and Classification: A Comprehensive Review. Journal of Fungi, 11(1), 758. <https://doi.org/10.3390/jof11010758>
- Rokh, B., Azarpeyvand, A., & Khanteymoori, A. (2023). A comprehensive survey on model quantization for deep neural networks in image classification. ACM Transactions on Intelligent Systems and Technology, 14(6), 1-50.
- Tan, M., & Le, Q. (2019, May). EfficientNet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning (pp. 6105-6114). PMLR.
- Tan, M., & Le, Q. (2021, July). Efficientnetv2: Smaller models and faster training. In International conference on machine learning (pp. 10096-10106). PMLR.

Pipeline



- Data Acquisition once in the beginning
- Done for different models:
 - Data Loading
 - Hyperparameter Optimization
 - Head Training
 - Fine-tuning (skipped for ResNet50)
 - Post Training Quantization (PTQ)
 - Evaluation

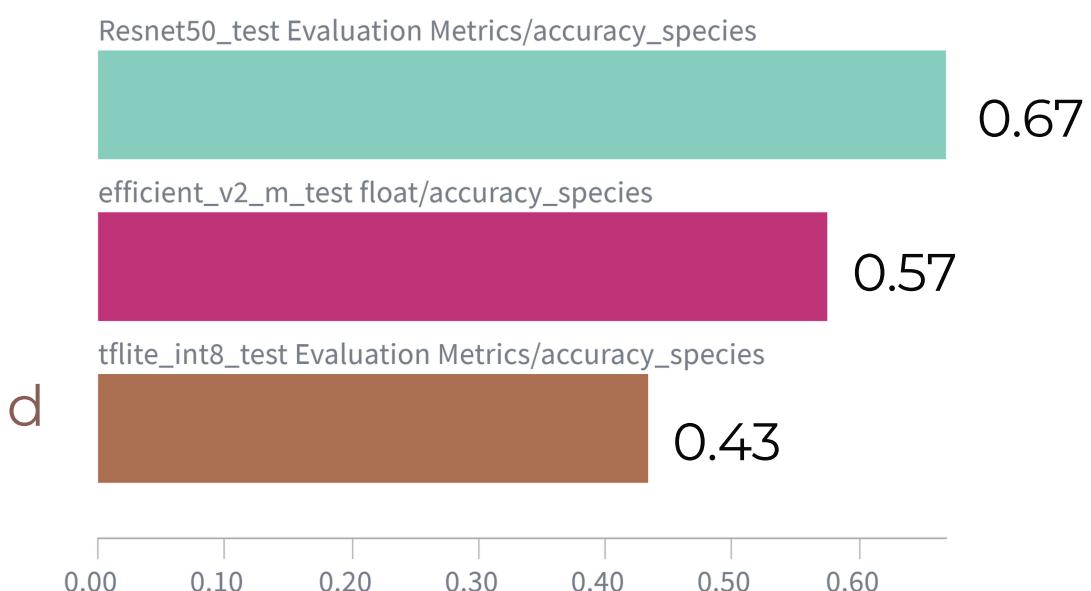
Results

Models

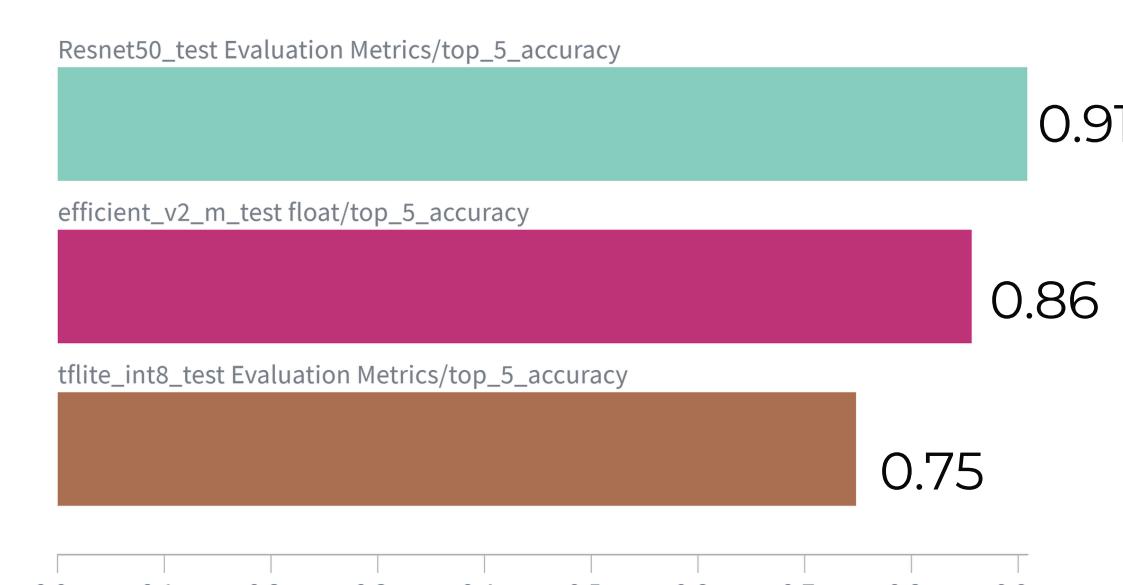
- Baseline: ResNet50
- EfficientNet-V2-M
- Quantized EfficientNet

- Baseline model achieved best results.
- Edibility precision too low for real world deployment.
- Finetuning regime and other models should be explored.

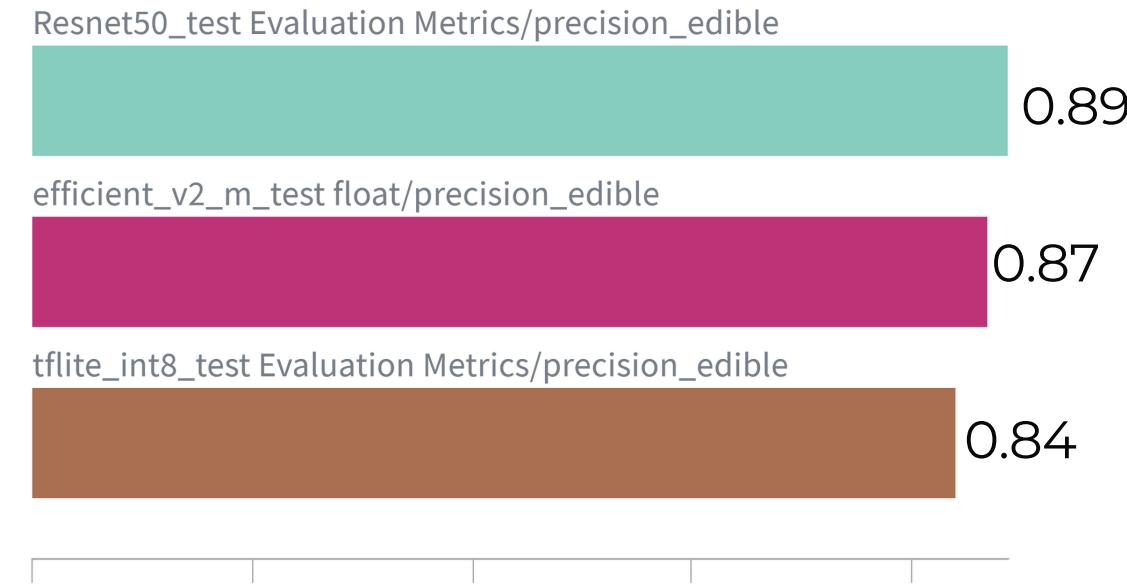
Species Accuracy



Species Top-5 Accuracy



Edibility Precision

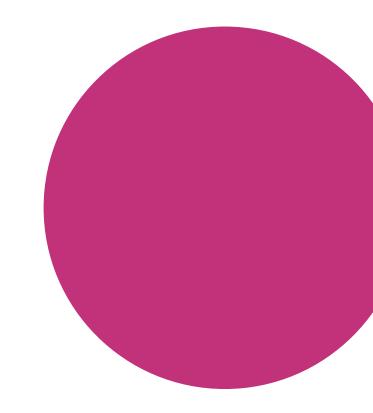


Compression

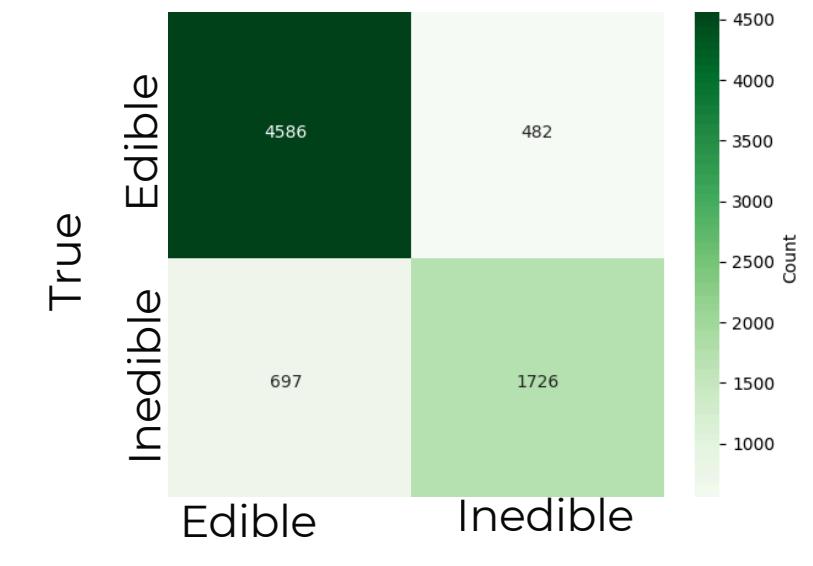
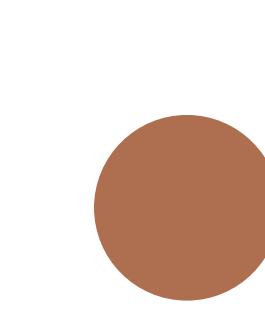
- Compression didn't affect species already misclassified, but introduced errors in additional species, driving most of the accuracy loss.
- model compression ratio = 0.4
- 0.1415 (-24.62%) drop in species accuracy
- 0.0302 (-3.47%) drop in edibility precision

Model Size

230 MB



57 MB

Edibility Confusion Matrix
Compressed EfficientNet-V2-M

Misclassified Mushrooms

Puffballs



Common Puffball



Pear Shaped Puffball

Boletes



Lurid Bolete



King Bolete

Panthercap



False Death Cap



Panthercap