
Investigating Target Class Influence on Neural Network Compressibility for Energy-Autonomous Avian Monitoring

Nina Brolich^{1,2,3}, Simon Geis¹, Maximilian Kasper¹,
Alexander Barnhill⁴, Axel Plinge¹, and Dominik Seuß^{1,5}

¹Fraunhofer Institute for Integrated Circuits IIS, Germany

²Fachhochschule Erfurt, Germany

³Universität Erfurt, Germany

⁴Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Germany

⁵Center for Artificial Intelligence and Robotics (CAIRO),

Technische Hochschule Würzburg-Schweinfurt, Germany

Abstract

Biodiversity loss poses a significant threat to humanity, making wildlife monitoring essential for assessing ecosystem health. Avian species are ideal subjects for this due to their popularity and the ease of identifying them through their distinctive songs. Traditional avian monitoring methods require manual counting and are therefore costly and inefficient. In passive acoustic monitoring, soundscapes are recorded over long periods of time. The recordings are analyzed to identify bird species afterwards. Machine learning methods have greatly expedited this process in a wide range of species and environments, however, existing solutions require complex models and substantial computational resources. Instead, we propose running machine learning models on inexpensive microcontrollers directly in the field. Due to the resulting hardware and energy constraints, efficient Artificial Intelligence (AI) architecture is required.

In this paper, we present our method for avian monitoring on microcontrollers. We trained and compressed models for various numbers of target

classes to assess the detection of multiple bird species on edge devices and evaluate the influence of the number of species on the compressibility of neural networks. Our results demonstrate significant compression rates with minimal performance loss. We also provide benchmarking results for different hardware platforms and evaluate the feasibility of deploying energy-autonomous devices.

Keywords: edge AI, biodiversity monitoring, energy-autonomous, energy benchmarking.

7.1 Introduction

Among today's global environmental crises, the ongoing loss of biodiversity stands out as especially severe: it ensures access to vital resources [1] and contributes to various environmental services such as climate regulation, control of pollution and soil erosion, as well as pollination of crops, while also holding intrinsic value for cultural identity [2]. Currently, biodiversity is declining at unprecedented rates, highlighting the need for conservation efforts [3].

In this context, biodiversity monitoring provides essential information for tracking environmental changes and plays an integral role in assessing population trajectories of different species [4]. Birds are particularly well suited for monitoring due to their frequent and distinctive vocalizations [5], their role as indicators of ecosystem health [6], and their popular appeal [7].

Avian monitoring has been done mainly through point counts [8], where the goal is to record all birds seen or heard within a given time period, with the observer being stationary [9]. However, point counts are highly susceptible to external circumstances, such as weather conditions, dependent on human expertise [10], and their scalability is limited by logistic and financial constraints [11].

In recent years, Autonomous Recording Units (ARUs) have emerged as a cost-effective alternative to point counts, allowing long-term sound collection (Passive Acoustic Monitoring (PAM)) with minimal disturbance, broader spatial and temporal coverage, and permanent recordings for re-analysis [12]. However, analyzing these large datasets is challenging [13]: manual methods require reducing the number of recordings or species (e.g. [14]), while automatic approaches remain difficult. Deep Neural Networks (DNNs) show promise but face computational constraints ([15, 16]). *BirdNET* [11], capable of identifying thousands of species, is an example of a successful DNN-based

solution and provides our method with a baseline for data collection and pre-processing. However, this approach is computationally expensive, requires internet access, and often exceeds the needs of localized surveys.

As an alternative, we propose species identification in real-time on an embedded, energy self-sustaining solution, shifting from retrospective analysis of ARU data to in-field processing, thereby reducing computational overhead while improving both energy and cost efficiency. The system is to be realized with edge Artificial Intelligence (AI) using a Microcontroller Unit (MCU), which introduces resource constraints such as limited memory, low processing capacity, and lack of parallelism [17]. To operate effectively under these constraints, deployment on edge devices like MCUs requires compact models, achieved either through lightweight network design or compression, to accommodate hardware constraints while preserving performance.

To assess the feasibility of avian monitoring on the edge and to examine target class influence on compressibility, we trained and compressed models for various numbers of avian target classes, benchmarked their energy consumption and latency on an *ARM Cortex-M4*, an *ARM Cortex-M7*, and a *Raspberry Pi 4*, and provided an estimate for the required battery capacity for real-life deployment. In the following, we describe our methodology, present and discuss the results, and conclude with an outlook on future work.

7.2 Methodology

An overview of our methodology is provided in Figure 7.1. The dataset was primarily compiled from *Xeno-Canto*, an open database of user-submitted bird recordings, resulting in a large but heterogeneous collection in both quality and duration. 500 species were selected based on the availability of sufficient samples, with preference first given to German, then European and finally global species. German species and data were prioritized to match the intended deployment region and to ensure data representativeness. For each species, 250 recordings were randomly chosen. In addition, the *ESC-50* [18] dataset was used to form a single non-avian class by merging 49 environmental sound categories (excluding bird calls). The data samples were pre-processed using the following steps:

1. Discarding audio samples shorter than 2 seconds, based on the average length of a bird vocalization of 1.94 seconds [19].
2. Removing silent sections, defined as amplitudes below 20 % of the maximum peak.

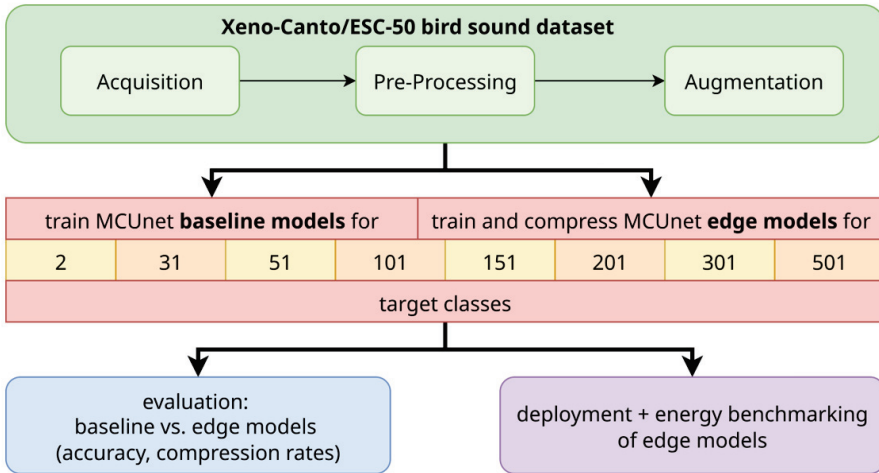


Figure 7.1 Methodology overview: workflow for dataset preparation, MCUNet model training, compression, and edge benchmarking

3. Sequentially splitting recordings into 2-second chunks with a maximum of 30 chunks per recording while discarding chunks without peaks (at least 7.5 % louder than surrounding points).
4. Normalizing each chunk so that its maximum absolute amplitude equals 1 to ensure consistency in amplitude throughout the dataset.
5. Converting the normalized chunks into mel-spectrograms, using a sample rate of 48 kHz, 64 mel bands, a Fast Fourier Transform (FFT) window size of 512 and a hop length of 384 [11]. The frequency was constrained to 150 Hz and 7.5 kHz.

This produced an easily attainable yet heterogeneous dataset, containing samples of varying quality with both avian and non-avian background noises. While not optimal for training, it realistically represents the acoustic challenges encountered in edge-deployed avian monitoring. On average, this resulted in 2452 chunks per bird species with a standard deviation of 874. The dataset was partitioned into training, validation, and test sets without data leakage.

To increase the diversity of our training data and to enhance the generalizability of our models, we applied the data augmentation pipeline outlined by [11] to our data. Four different augmentation methods were applied, as illustrated in Figure 7.2:

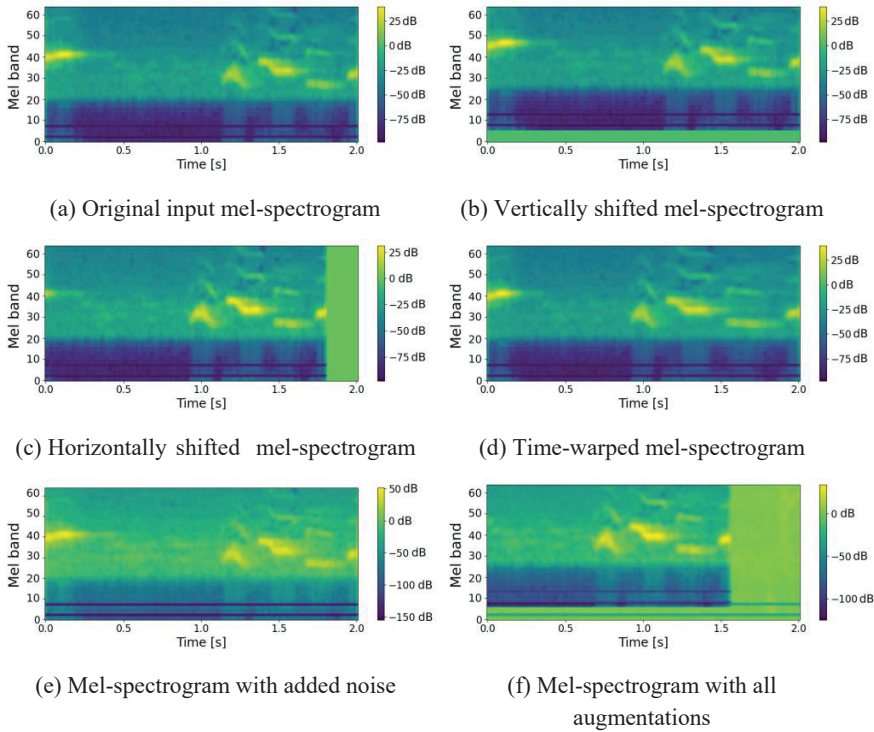


Figure 7.2 Mel-spectrograms with different data augmentation methods applied

- **Random vertical shifts (frequency roll):** the mel-spectrogram was shifted up or down by a random factor between -5 % and 5 %, with the vacated area filled with zeros.
- **Random horizontal shifts (time roll):** the mel-spectrogram was shifted left or right by a random factor between -25 % and 25 %, with zeros filling the empty region.
- **Time warping:** the mel-spectrogram was deformed in time direction using the *SpecAugment* algorithm [20].
- **Addition of noise:** randomly selected noise chunks, previously discarded during preprocessing and verified to contain no bird calls, were added at a random intensity between 20 % and 80 %.

These augmentation methods all represent acoustic variations between training and test data like the changes in the vocal output of birds depending on environmental factors, high levels of ambient noise, and lack of training sample diversity [11]. Each augmentation was applied with a 50 % probability

per chunk, in random order, with a maximum of three augmentations per chunk.

For the experiments, the *mcunet-in4* model from the *MCUNet* framework [21] was selected and adapted for bird sound classification. *MCUNet* is specifically designed for deep learning on microcontrollers, combining efficient neural architecture search (*TinyNAS*) with a memory-optimized inference engine (*TinyEngine*) to accommodate hardware constraints. Despite the existence of smaller *MCUNet* models, *mcunet-in4* was chosen for its better performance and to ensure support for higher numbers of target classes. Nevertheless, it still allows deployment on edge devices, which was also aided by further compression of the models. Its architecture, illustrated in Figure 7.3, consists of an initial convolutional layer, 17 *MobileInvertedResidualBlocks* [22], and a final linear layer. To accommodate mel-spectrogram inputs, the first layer was modified to accept a single input channel, while the output layer was adjusted to match the number of target classes (31, 51, 101, etc.), including a non-avian class. Pre-trained weights from ImageNet were loaded for all layers except the first and last.

In total, one uncompressed baseline model and 50 compressed edge models were trained for 2, 31, 51, 101, 151, 201, 301, and 501 target classes. For the model with only two target classes, the objective was to identify one bird species against 29 others, as well as non-bird data, i.e., it was trained with two classes, one consisting of data for the common blackbird (*turdus merula*), and the other one consisting of data of the 29 other most common

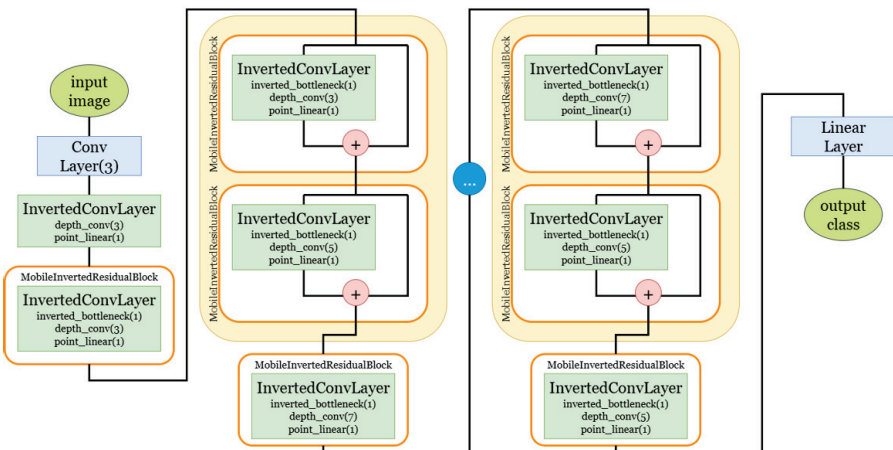


Figure 7.3 Structure and components of *MCUNet*.

bird species in Germany, as well as the environment data. From 31 target classes onward, the models were trained to identify 30 (or 50, 100, ...) bird species, as well as one non-event class.

The baseline models were trained with a set number of 30 training epochs. The training was conducted with a batch size of 32, using the Stochastic Gradient Descent (SGD) optimizer with a learning rate of 0.001 and a momentum of 0.9. Cross-entropy loss served as the loss function. The edge models were trained and compressed using an internal tool from *Fraunhofer IIS*. While the same general training configuration was utilized as for the baseline models, an interleaved pruning process with quantization at the end of the training process was employed. The tool implements the neural network compression approach from [23], performing multiple training and compression trials to maximize accuracy and minimize Read-Only Memory (ROM), Random-Access Memory (RAM), and Floating Point Operations (FLOPs) requirements. For each edge model, 50 trials were run, producing 50 compressed edge models, some of which are Pareto optimal across these objectives. A model is considered Pareto optimal if no other model performs better in one objective without performing worse in at least one other.

For the evaluation of the results, the performance of the edge models was first compared to the corresponding baseline models. A representative compressed model was selected from the 50 trials conducted for each target class number. To rank the trial t_x , the trade-off between accuracy (Eq. (7.1)) and the combined metrics of ROM, RAM, and FLOPs (Eq. (7.2)) was emphasized, as showcased in the ranking function r in Eq. (7.3).

$$\text{acc}(t_x) = \frac{\text{ACC}(t_x)}{\max\{\text{ACC}(t_i) \mid i \in [0, 49]\}} \quad (7.1)$$

$$\text{mem}(t_x) = \frac{1 - \frac{\text{RAM}(t_x)}{\max\{\text{RAM}(t_i)\}} + 1 - \frac{\text{ROM}(t_x)}{\max\{\text{ROM}(t_i)\}} + 1 - \frac{\text{FLOPS}(t_x)}{\max\{\text{FLOPS}(t_i)\}}}{3}, i \in [0, 49] \quad (7.2)$$

$$r(t_x) = \text{mem}(t_x) + \text{acc}(t_x) \quad (7.3)$$

To evaluate compressibility, we defined ROM, RAM, and FLOPs compression rates cr , shown in Eq. (7.4) for FLOPs, with the values for ROM and RAM computed analogously.

$$cr_{\text{FLOPs}}(\text{FLOPs}_{\text{baseline}}, \text{FLOPs}_{\text{edge}}) = 1 - \frac{\text{FLOPs}_{\text{edge}}}{\text{FLOPs}_{\text{baseline}}} \quad (7.4)$$

The overall compression rate of a model was defined as the mean of its FLOPs, ROM, and RAM compression rates, while the average overall compression rate was computed as the mean of these values across all Pareto-optimal trials for the target classes.

The *dnnruntime* framework [24] was used to convert the models to C code and create a deployable binary file. This binary file was then flashed onto *ARM Cortex-M4* and *ARM Cortex-M7* processors on a *SparkFun Micro-Mod ATP* carrier board to measure the latency and energy consumption of the models on the two microcontrollers. Additionally, we benchmarked the models on a *Raspberry Pi 4B* using the *onnxruntime* framework [25]. For an estimation of the feasibility of energy-autonomous deployment, we make the following assumptions:

1. The device is equipped with a battery which should be able to power the device for at least 48 hours without charging.
2. In addition, the device is equipped with a solar panel that should be able to fully charge the battery in 24 hours.
3. The device is in sleep mode per default and wakes up every 10 seconds. If the device recognizes sound from the microphone, inference is performed as long as sound is detected. Otherwise, the device returns to sleep for the next 10 seconds. We assume that the inference step is performed 10 % of the time.

Based on the power required during inference and idle, we can infer the required battery size of the device for running for 48 hours. The required size A of the solar cell can then be estimated by the average sun-radiation power density S_{rad} in Germany and by the required charging output P_{charge} as shown in Eq. (7.5). We assume a solar panel efficiency η_{solar} of 20 % and a charging efficiency η_{bat} of 90 %.

$$A = \frac{P_{charge}}{\eta_{solar} \cdot \eta_{solar} \cdot S_{rad}} \quad (7.5)$$

7.3 Results

To assess the general performance of the models, one baseline and one representative edge model per target class number were selected according to the rank defined in Eq. (7.3). Their validation accuracies are shown in Figure 7.4.

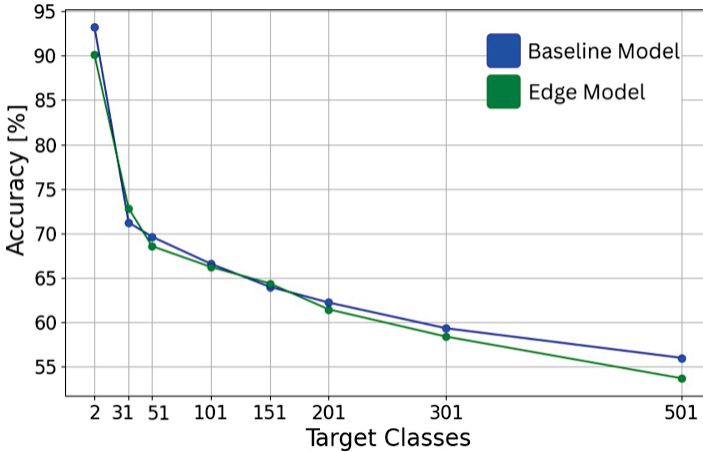


Figure 7.4 Validation accuracies for one baseline and one edge model per number of target classes.

Overall, validation accuracy declines as the number of target classes increases. While the models achieve high accuracy values for lower target class numbers, performance is more moderate with more target classes. Notably, there is almost no accuracy loss due to compression, as the accuracies of the baseline and edge models remain within a very similar range. To assess target class influence on compressibility, the average compression rate was computed across different target class numbers and is shown in Figure 7.5.

The results indicate a slight decline in compressibility with increasing class numbers. From 201 classes onward, however, this trend reverses, with compressibility improving for larger class counts. This result is unexpected, as a larger number of target classes would hypothetically require more complex architectures with higher memory and FLOPs demands, thereby reducing compressibility. While this trend is observed up to a certain point, it appears to reverse for larger class counts. Since the training and compression procedures are essentially a black box, a definitive explanation is difficult. Nonetheless, several factors may contribute: with more classes, models may exploit feature sharing more effectively, capturing overlapping features with fewer parameters [26]. Moreover, the inclusion of additional classes may encourage the learning of more generalized and compact representations, ultimately reducing resource requirements [27]. However, the observed decrease and subsequent increase in compressibility are subtle, with compression rates

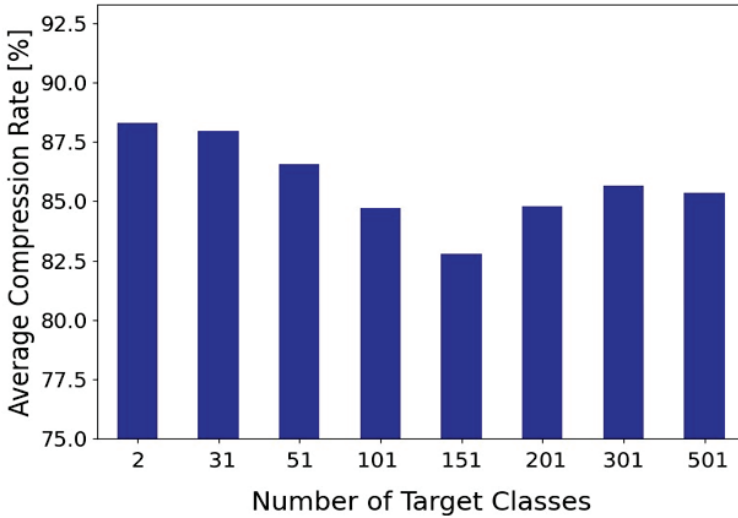


Figure 7.5 Average overall compression rate for all Pareto optimal trials per number of target classes regarding the reduction in RAM, ROM and FLOPs

ranging only between approximately 82 % and 88 % and thus may not reflect a clear or consistent trend.

To evaluate the feasibility of deploying the compressed models on an MCU, energy consumption and latency were measured for one model per number of target classes, selected according to the ranking in Eq. (7.3). The results are shown in Figure 7.6.

Overall, both latency and energy consumption increase with a growing number of target classes, though the improved compressibility of models with more than 151 classes is partially reflected in lower values for these metrics. Because both energy consumption and latency are largely determined by the number of FLOPs performed during inference, they are only indirectly related to compressibility. Since model selection accounted for FLOPs, ROM, and RAM, cases arise where a model with more FLOPs than its successor was chosen due to lower memory requirements. This explains, for example, why the 31-class model consumes more energy than the 51-class model. Finally, the benchmarking results indicate that the compressed models achieve energy and latency values suitable for real-world deployment on the *ARM Cortex-M7* and *Raspberry Pi 4*. In contrast, on the more resource-constrained *ARM Cortex-M4*, latency consistently exceeded the audio chunk length, making real-world deployment infeasible.

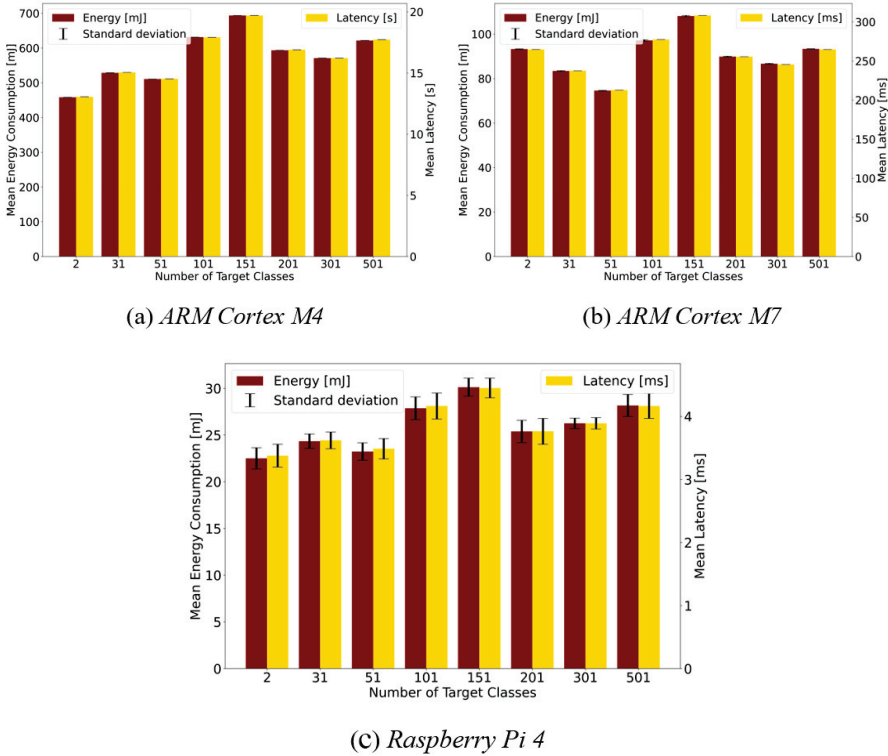


Figure 7.6 Average energy consumption and latency of one inference step of the best ranked model for each number of target classes

For the evaluation of energy-autonomous devices, we selected the model with 31 classes as an example. We measured an energy consumption of 83 mJ, a latency of 237 ms for the inference, and 55 mJ and 170 ms for the spectrogram generation on the *ARM Cortex M7*. This results in an average power of 339 mW. During sleep, we measured a power consumption of 116 mW, resulting in an average power consumption of 138.3 mW during the day. Overall, this leads to a required battery capacity of 6.6 Wh.

For the *Raspberry Pi 4*, we measured 24.3 mJ and 3.6 ms for the inference and 483 mJ and 80.9 ms for the mel-spectrogram generation, leading to an average power of 6.0 W. During idle, we measured a power consumption of 2.93 W, resulting in an average power of 3.24 W. In total, this would result in a required battery capacity of 155.5 Wh.

Based on these values, the required size of the solar panel can be calculated as described in Eq. (7.5). During December, the sun has the lowest

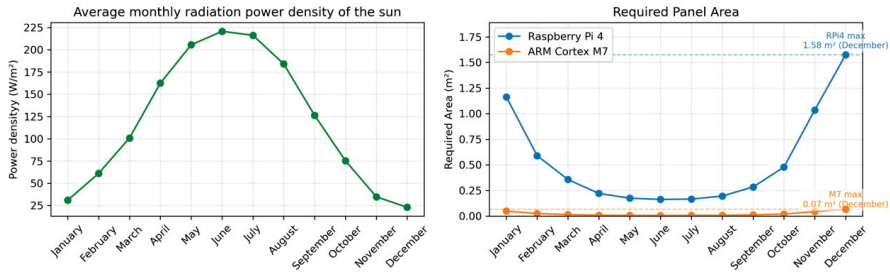


Figure 7.7 Average power density of the sun’s radiation in Germany [28] (left) and the resulting area of the solar panel for each month (right).

power density of 22.8 Wm^{-2} as shown in Figure 7.7. With a required charging power of 275 mW, this results in a required panel area of 0.07 m^2 for the *ARM Cortex M7*. The *Raspberry Pi* requires a charging power of 6.48 W and therefore a panel size of 1.58 m^2 .

7.4 Conclusion

Avian species identification is a challenging task, with model accuracy generally decreasing as the number of target classes increases. This is partly due to the heterogeneity of the dataset and the use of a single, relatively simple model architecture across all class numbers, which reflects realistic conditions for edge deployment and was necessary to facilitate comparability of compressibility across different target class numbers.

Our results demonstrate that high compression rates are achievable with minimal loss of accuracy across different numbers of target classes. Although compressibility initially decreased with increasing class numbers and later increased beyond 151 classes, the overall variation is minor and does not necessarily suggest a clear trend. Instead, it reflects the complex interplay between task complexity, model architecture, and learned representations. The evaluation of energy consumption and latency shows that real-world deployment is feasible on the *Raspberry Pi 4* and on the *ARM Cortex-M7*, but not on the more constrained *ARM Cortex-M4*, emphasizing the importance of selecting appropriate hardware for edge applications.

Overall, we showed that neural network compression is a practical and effective strategy for edge-based avian monitoring, even for large numbers of target classes. Finally, our results indicate that avian monitoring is feasible on energy-autonomous edge devices which could play a crucial role in wildlife monitoring and biodiversity conservation.

Future work could include curating a scientific dataset, exploring different species arrangements in the training data, comparing alternative compression frameworks, and investigating end-to-end audio models to reduce pre-processing overhead. Furthermore, real-life deployment will require additional functionalities, such as counting, data storage, and energy management, potentially combining edge AI with automated data collection technologies.

Acknowledgement

This work is part of the GreenICT@FMD project and is funded by the German Federal Ministry for Research, Technology and Space (BMFTR) (grant number 16ME0491K).

References

- [1] V. Singh, S. Shukla, and A. Singh, “The principal factors responsible for biodiversity loss,” *Open Journal of Plant Science*, vol. 6, no. 1, pp. 11–14, 2021. <https://www.agriscigroup.us/articles/OJPS-6-126.php>
- [2] M. S. Habibullah, B. H. Din, S.-H. Tan, and H. Zahid, “Impact of climate change on biodiversity loss: global evidence,” *Environmental Science and Pollution Research*, vol. 29, no. 1, pp. 1073–1086, 2022. <https://doi.org/10.1007/s11356-021-15702-8>
- [3] R. E. Almond, M. Grooten, and T. Peterson, *Living Planet Report 2020—Bending the Curve of Biodiversity Loss*. World Wildlife Fund, 2020. <https://pure.iiasa.ac.at/id/eprint/16870/1/ENGLISH-FULL.pdf>
- [4] D. Schmeller, K. Henle, A. Loyau, A. Besnard, and P.-Y. Henry, “Bird monitoring in Europe—A first overview of practices, motivations and aims,” *Nature Conservation*, vol. 2, pp. 41–57, 2012. <https://doi.org/10.3897/natureconservation.2.3644>
- [5] J. Shonfield and E. M. Bayne, “Autonomous recording units in avian ecological research: current use and future applications,” *Avian Conservation & Ecology*, vol. 12, no. 1, 2017. <https://doi.org/10.5751/ACE-00974-120114>
- [6] S. Mekonen, “Birds as biodiversity and environmental indicator,” *Journal of Natural Sciences Research*, vol. 7, no. 21, 2017. <https://www.iiste.org/Journals/index.php/JNSR/article/view/39931/41049>

- [7] H. Şekercioğlu, “Promoting community-based bird monitoring in the tropics: conservation, research, environmental education, capacity-building, and local incomes,” *Biological Conservation*, vol. 151, no. 1, pp. 69–73, 2012. <https://doi.org/10.1016/j.biocon.2011.10.024>
- [8] V. Cavarzere, G. P. Moraes, J. J. Roper, L. F. Silveira, and R. J. Donatelli, “Recommendations for monitoring avian populations with point counts: a case study in southeastern Brazil,” *Papéis Avulsos de Zoologia*, vol. 53, pp. 439–449, 2013. <https://doi.org/10.1590/S0031-10492013003200001>
- [9] M. A. Etterson, G. J. Niemi, and N. P. Danz, “Estimating the effects of detection heterogeneity and overdispersion on trends estimated from avian point counts,” *Ecological Applications*, vol. 19, no. 8, pp. 2049–2066, 2009. <https://doi.org/10.1890/08-1317.1>
- [10] J. P. Brewster and T. R. Simons, “Testing the importance of auditory detections in avian point counts,” *Journal of Field Ornithology*, vol. 80, no. 2, pp. 178–182, 2009. <https://doi.org/10.1111/j.1557-9263.2009.00220.x>
- [11] S. Kahl, C. M. Wood, M. Eibl, and H. Klinck, “BirdNET: A deep learning solution for avian diversity monitoring,” *Ecological Informatics*, vol. 61, p. 101236, 2021. <https://doi.org/10.1016/j.ecoinf.2021.101236>
- [12] C. Pérez-Granados and J. Traba, “Estimating bird density using passive acoustic monitoring: a review of methods and suggestions for further research,” *Ibis*, vol. 163, no. 3, pp. 765–783, 2021, <https://doi.org/10.1111/ibi.12944>
- [13] I. Molina-Mora, V. Ruíz-Gutiérrez, M. Vega-Hidalgo, and L. Sandoval, “The utility of passive acoustic monitoring for using birds as indicators of sustainable agricultural management practices,” *Frontiers in Bird Science*, vol. 3, p. 1386759, 2024. <https://doi.org/10.3389/fbirs.2024.1386759>
- [14] B. J. Furnas, “Rapid and varied responses of songbirds to climate change in California coniferous forests,” *Biological Conservation*, vol. 241, p. 108347, 2020. <https://doi.org/10.1016/j.biocon.2019.108347>
- [15] I. Potamitis, S. Ntalampiras, O. Jahn, and K. Riede, “Automatic bird sound detection in long real-field recordings: applications and tools,” *Applied Acoustics*, vol. 80, pp. 1–9, 2014. <https://doi.org/10.1016/j.apacoust.2014.01.001>
- [16] E. Znidarsic, M. Towsey, W. K. Roy, S. E. Darling, A. Truskinger, P. Roe, and D. M. Watson, “Using visualization and machine learning methods to monitor low-detectability species—The least bittern as a

- case study,” *Ecological Informatics*, vol. 55, p. 101014, 2020. <https://doi.org/10.1016/j.ecoinf.2019.101014>
- [17] B. Sudharsan, J. G. Breslin, and M. I. Ali, “ML-MCU: A framework to train ML classifiers on MCU-based IoT edge devices,” *IEEE Internet of Things Journal*, vol. 9, no. 16, pp. 15007–15017, 2021. <https://doi.org/10.1109/JIOT.2021.3098166>
- [18] K. J. Piczak, “ESC: Dataset for environmental sound classification,” in *Proceedings of the 23rd ACM International Conference on Multimedia*, 2015, pp. 1015–1018. <https://doi.org/10.1145/2733373.2806390>
- [19] S. Kahl, “Identifying birds by sound: Large-scale acoustic event recognition for avian activity monitoring,” Ph.D. dissertation, Chemnitz Univ. of Technology, 2019. <https://nbn-resolving.org/urn:nbn:de:bsz:ch1-qucosa2-369869>
- [20] D. S. Park, W. Chan, Y. Zhang, C.-C. Chiu, B. Zoph, E. D. Cubuk, and Q. V. Le, “SpecAugment: A simple data augmentation method for automatic speech recognition,” *arXiv preprint arXiv:1904.08779*, 2019. <https://doi.org/10.48550/arXiv.1904.08779>
- [21] J. Lin, W.-M. Chen, Y. Lin, C. Gan, S. Han, *et al.*, “MCUNet: Tiny deep learning on IoT devices,” *Advances in Neural Information Processing Systems*, vol. 33, pp. 11711–11722, 2020. https://proceedings.neurips.cc/paper_files/paper/2020/file/86c51678350f656dcc7f490a43946ee5-Paper.pdf
- [22] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “MobileNetV2: Inverted residuals and linear bottlenecks,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 4510–4520. https://openaccess.thecvf.com/content_cvpr_2018/html/Sandler_MobileNetV2_Inverted_Residuals_CVPR_2018_paper.html
- [23] M. Deutel, G. Kontes, C. Mutschler, and J. Teich, “Combining multi-objective Bayesian optimization with reinforcement learning for TinyML,” *ACM Transactions on Evolutionary Learning*, vol. 5, no. 3, pp. 1–21, 2025. <https://doi.org/10.1145/3715012>
- [24] M. Deutel, P. Woller, C. Mutschler, and J. Teich, “Energy-efficient deployment of deep learning applications on Cortex-M based micro-controllers using deep compression,” in *MBMV 2023: 26th Workshop, VDE*, 2023, pp. 1–12. <https://ieeexplore.ieee.org/document/10173060>
- [25] ONNX Runtime Developers, “ONNX runtime,” 2021. <https://onnxruntime.ai/>
- [26] G. Liu, H. Zeng, and D. K. Gifford, “Visualizing complex feature interactions and feature sharing in genomic deep neural networks,” *BMC*

Bioinformatics, vol. 20, pp. 1–14, 2019. <https://doi.org/10.1186/s12859-019-2957-4>

- [27] F. Zhu, X.-Y. Zhang, R.-Q. Wang, and C.-L. Liu, “Learning by seeing more classes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 6, pp. 7477–7493, 2022. <https://doi.org/10.1109/TPAMI.2022.3225117>
- [28] Deutscher Wetterdienst, “Global radiation in Germany 1991 - 2020,” Available: https://www.dwd.de/EN/ourservices/solarenergy/maps_globalradiation_mvs.html.