Evaluation of Deep Learning Models that Recognize Emotions for IoT systems

Tanvi S. Johari (20BCE0619)

Dipsikha De (20BCE0631)

Kaushik S. Prabhakar (20BCE0635)

Harikrishnan Nair (20BCE0668)

A report submitted for the J component of

CSE3009 - Internet of Things

under the guidance of Dr. Dheeba J

School of Computer Science and Engineering Vellore institute of Technology, Vellore

Table of Contents

Introduction

Aim

The aim of this project is to evaluate deep learning models for mood detection based on facial recognition and image processing stages like capturing images, pre-processing them, and then feeding them to the models, against self defined metrics to make facial recognition more efficient in IoT applications.

Objective

Most deep learning applications are only quantified with respect to its accuracy in making predictions or classifications. In an IoT application, this is not the only aspect of concern. Through this project we wish to quantise metrics that are usually ignored that are essential in an IoT application. These metrics can serve as a benchmark when developing deep learning models that are to be deployed in IoT applications.

Benefits

IoT applications involve the use of embedded systems and other targeted hardware/software components. The integration of facial recognition involves the use of sensors like cameras. Deep learning specialists can optimize the power draw, time taken, battery usage, etc by sensors as such to increase longevity of IoT devices.

Literature Survey

S.No.	Name of the Paper	Author and Year	Major Technologies	Result Outcome	Drawbacks
1	Mood based music recom-mendation system	Ankita Mahadik, Shambhavi Milgir, Janvi Patel, Prof. Vijaya Bharati Jagan & Prof. Vaishali Kavathekar *March 2022*	FaceDetector , MobileNet (CNN architecture)	MobileNet required very less computation power to run, and yielded an accuracy of 75% for detection of emotions.	To accurately detect moods such as fear and disgust, additional parameters such as heart rate / body temperature must also be considered.
2	Smart music player integrating facial emotion recognition and music mood recommendation	Shlok Gilda, Husain Zafar, Chintan Soni, Kshitija Waghurdekar March 2017	Recommender systems, Artificial neural networks, Multi-layer neural network	The Emotion Module identifies the user's mood with an accuracy of 90.23%. The Music classification Module achieves a remarkable result of 97.69% while classifying songs into 4 different mood classes.	Despite impressive results, there is room for improvement. Moreover, additional songs from different languages and regions can also be added to make the recommendation system more robust. Overall system can be improved using collaborative filtering.
3	Music Recommend-ation System Based On Facial Emotion Recognition	Deny John Samuvel, B. Perumal, Muthukumaran Elangovan	Local Binary Patterns, Direct Cosines Transform, Gabor Wavelets, EigenFaces, SVM classifier,		Progressively effective approaches should be investigated due to the lopsided nature of each element set. The informational collection used to

			ANN		construct the grouping model could be expanded further to improve the accuracy of the arrangement framework.
4	Facial Expression Based Music Recommend-ation System	Vinay P, Raj Prabhu T, Bhargav Satish Kumar Y, Jayanth P, A. Suneetha	Tkinter module, OpenCV, SVM Algorithm	The accuracy of the emotion detection algorithm for real time images is around 85-90%, while for static images it is around 98- 100%.	System relies on manual image upload by user instead of real-time facial expression recognition using a camera module
5	Music Recommendation Based on Face Emotion Recognition	Madhuri Athavle, Deepali Mudale, Upasana Shrivastav, Megha Gupta	LBP (Local Binary Pattern) and Haar Cascades, Haar Wavelet , Pygame, Tkinter, Convolutional Neural Networks,	Testing accuracy for SVM, ELM and CNN is 66%, 63%, 71% respectively.	The current system does not perform well in extremely bad light conditions and poor camera resolution thereby provides an opportunity to add some functionality as a solution in the future.
6	Music Recommendation based on Facial Expression using Deep Learning	Tanushree Gorasiya, Anushka Gore, Dimple Ingale, Megha Trivedi	Haar Cascade, MaxPooling, CNNs		Existing system struggles in extremely low light and has limited cam quality; however, certain capabilities could be improved as a solution
7	Song Recommendation System	Mohini , Aditi Singh, Afreen Khan	OpenCV, K-Nearest Neighbors	Results showed that 78% of end users preferred to	There is ambiguity with respect to the accuracy of the

	Using Facial Expression			listen to a musical genre similar to their current sentimental state, and only 22% preferred to listen to a different musical genre in relation to their current sentimental state.	system. There is also no mention of system performance in extreme conditions.
8	Face Detection and Verification Using Lensless Cameras	Tan, J. Niu, L. Adams, J.K. Boominathan, V. Robinson, J.T. Baraniuk, R.G. Veeraraghavan, A. December 2018	Faster R-CNN, VGG	Training on standard images yields poor accuracy while training on display-captured lensless images yields much improved accuracy. The FCFD Dataset had the most accurate results.	The performance gap between lensless and lens-based inference needs to be reduced. There are cost and size constraints while performing face detection and verification tasks.
9	Using deep learning approach and IoT architecture to build the intelligent music recommendation system	Xinglin Wen October 2020	Scale - Invariant Feature Transform, SVM, Fast-RCNN	The algorithm has a better performance in anti-interference capacity, robustness, and recognition capacity.	More compatibility tests and fault tolerance tests should be conducted.
10	LSTM-Based Emotion Detection Using Physiological	Muhammad Awais, Mohsin Raza, Nishant Singh, Kiran	-, Artificial Neural Networks - Long short-Term	LSTM models performed comparatively well by achieving performance of	The paper can be further extended with focus on end-to-end communications and

	Signals: IoT Framework for Healthcare and Distance Learning in COVID-19	Bashir, Umar Manzoor, Saif ul Islam, Joel J. P. C. Rodrigues	Memory (LSTM)	above 90% in classifying each of the four emotions. The worst performer for each class was a combination of only EMG sensors (C1) whose performance was around 70% for most of the classes.	visual aids to support distance learning. In addition, the incorporation of edge services can enhance the feasibility of the proposed work.
11	A Virtual Emotion Detection Architecture With Two-Way Enabled Delay Bound toward Evolutional Emotion-Based IoT Services	Kim, H. Ben-Othman, J. Mokdad, L. Bellavista, P.	Two-Way-Enabled-Border- Slab April 2022	A novel Two Way Enabled Border Slab scheme was proposed based on Emotion Services System Initialization. It was analyzed through extensive simulations.	Future works include: 1. Various target areas for virtual emotion detection 2. Applicability of Location Services 3. Security and Privacy of Virtual Emotion Systems
12	Emotion Detection in IoT-Based E-Learning Using Convolution Neural Network	Latha Parthiban and S. Selvakumara Samy	local binary patterns, fuzzy convolutional neural network (FCNN)	Face recognition across feature descriptors and the PCA & CNN combination gave accuracy levels of 86.9, 86.7 and 82.5% for FBET, SLE and eNTERFACE respectively	Future work includes further improvement of testing accuracy of the fuzzy CNN.

13	Federated Learning Meets Human Emotions: A Decentralized Framework for Human–Computer Interaction for IoT Application	Prateek Chhikara , Prabhjot Singh, Rajkumar Tekchandani , Neeraj Kumar , and Mohsen Guizani	federated learning, machine learning, sentiment analysis.	The proposed facial and speech emotion recognition classifier gives an accuracy of 71.64% and 85.04%, respectively	-
14	IoT based Real-time Face Detection and Recognition System	K. Kapoor, K. Gupta, N. Rakesh, R. Gusain, M. Kaur and P. Nand March 2022	Raspberry Pi, Haar Cascade Algorithm, OpenCV	The model has high accuracy and low latency.	Mask Detection Capabilities can be integrated into the system.
15	Deep Learning based Emotion Recognition IoT System	Kentaro YOKOO, Masahiko ATSUMI, Kei TANAKA, Haoqing WANG and Lin MENG	Raspberry Pi, Deep learning, Mobilenet	Experimental results show the motion sensor and RGB camera work well, and human identification almost achieves 100%, and emotion recognition also performs optimally.	Future work includes improving the emotion recognition rate and adding the subjects of the experiments for realizing the productization.
16	Optimized Face Detection and Alignment for Low-Cost and Low-Power IoT Systems	K. Choi and G. E. Sobelman	MTCNN (Multi Task CNN), FPGA System	This model has a 2.67 times lower power consumption. Model has 15.2 FPS, 54 GOPS, 0.28 J/Frame.	The model can be further optimized in terms of FPS and power consumption.
17	Facial	C. Le and T. K.	Mask RCNN,	71% of more	The algorithm can

	Detection in Low Light Environments Using OpenCV	Mohd	Faster RCNN, OpenCV	objects were detected using this algorithm. The average precision and recall were 0.81 and 0.79 respectively. The algorithm took 2s/image.	be further optimized. It is needed for the input image to have natural lighting with natural colour.
18	Key facial points recognition using ResNet	Swastik Kumar Sahu, Ram Narayan Yadav	ResNet, Key Facial Points	The accuracy of the ResNet model was 77.70% at the end of 100 epochs.	This ResNet model can be further improved by changing the model architecture like adding more RES blocks.
19	Emotion recognition from scrambled facial images via many graph embedding	Richard Jiang, AnthonyT. S. Ho, Ismahane Cheheb, Noor Al-Maadeed, Somaya Al-Maadeed and Ahmed Bouridane	Chaotic patterns Many graph embedding	proposed MGE method consistently attained the best average accuracy over all facial expressions	-
20	Extended deep neural network for facial emotion recognition	Deepak Kumar Jaina , Pourya Shamsolmoali b , Paramjit Sehdevc	Uncanny Valley, Embedded System, Emotional Intelligence	Proposed model has 0.32%, 0.34% better performance as compared to Jain et al. [1] and Zhang et al. [24] respectively.	For future work, the combination of FCN and residual block could considerably improve the overall result, which verified the efficiency of the proposed model.

Table 1: Literature Survey

Proposed Architecture Model

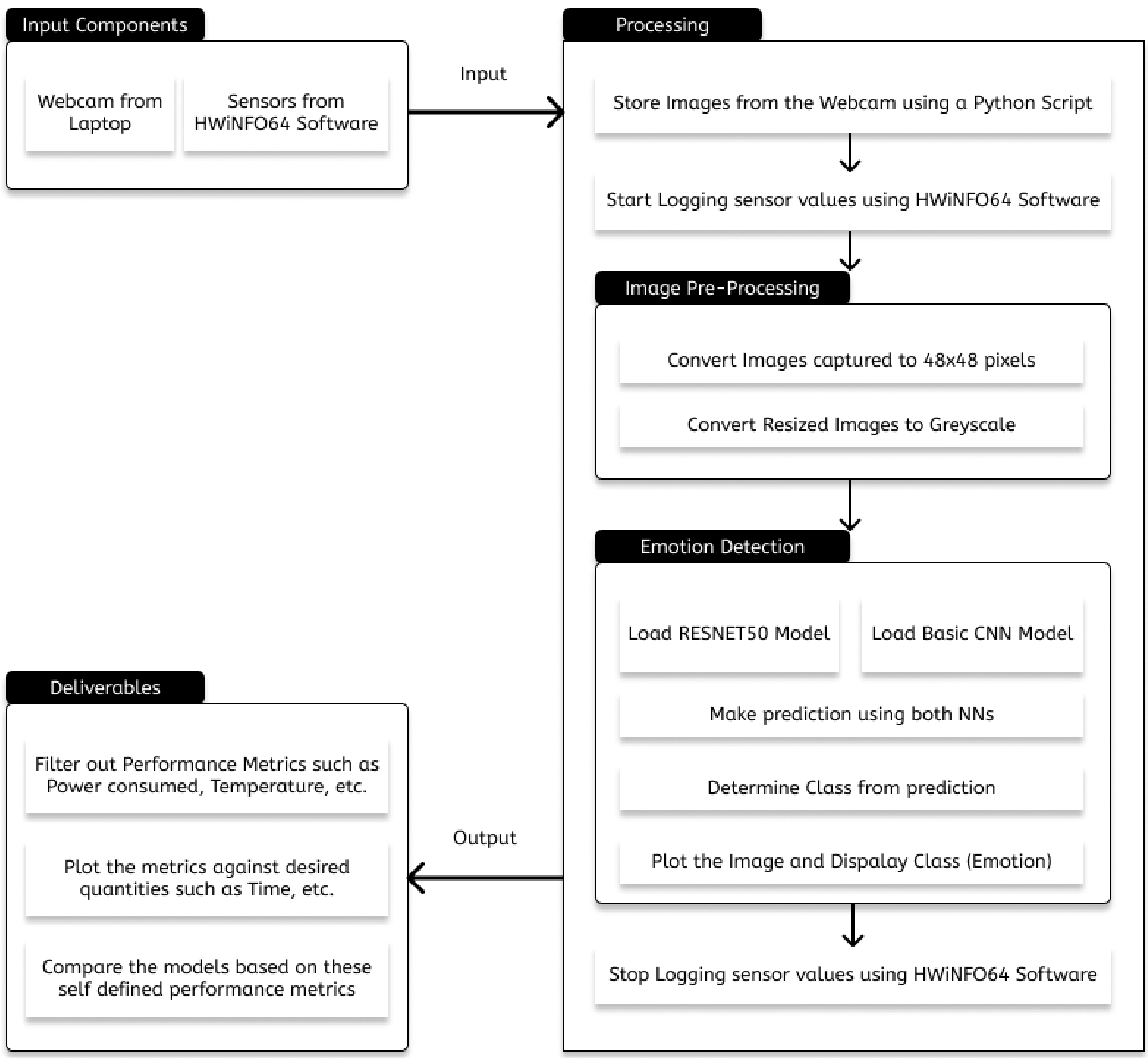

Figure 1: Architectural Model

To measure performance metrics before deploying a deep learning model in an IoT application we propose the following system as depicted in the Architectural Model shown above. The process can be broken down to 3 phases:-

Phase 1: Identifying input components. In this scenario of deep learning application, the inputs were an image taken by the camera. The internal sensors of a laptop provide us with details regarding performance which was made available to us using a software application HWInfo64 v7.32.

Phase 2: In this phase, the images are captured using a python script and stored to a directory. We then begin logging the internal sensors using HWInfo64 and then run a stress test on the two models where the image is preprocessed and classified for a large number of times.

This is done for the CNN Model and the RESNET50 seperately. We then stop logging the internal sensor readings namely Core Temperature(°C) ,Charge Level (%), Remaining Capacity (Wh) and CPU Package Power (W) and proceed to the next phase of our model.

Phase 3: In the last phase, we get the deliverables from the system which includes various parameters measured as mentioned in Phase 2, assessing the two models, and present a comparison between them using various graphs and charts.

Results and Discussion

The data from HWInfo64 was obtained in the form of a CSV file. The data was plotted against time in seconds and the following graphs were observed.

DEVICE 1

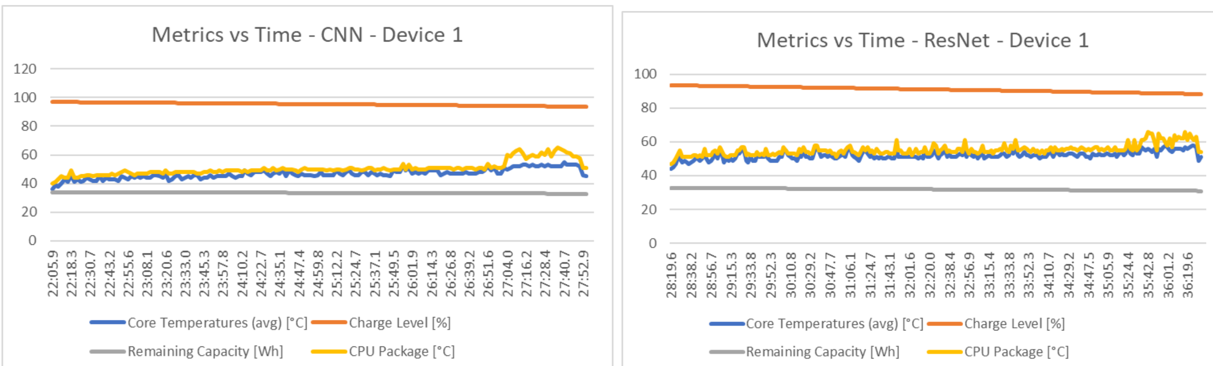

Figure 2: Visualising all Metrics against Time for Device 1

DEVICE 2

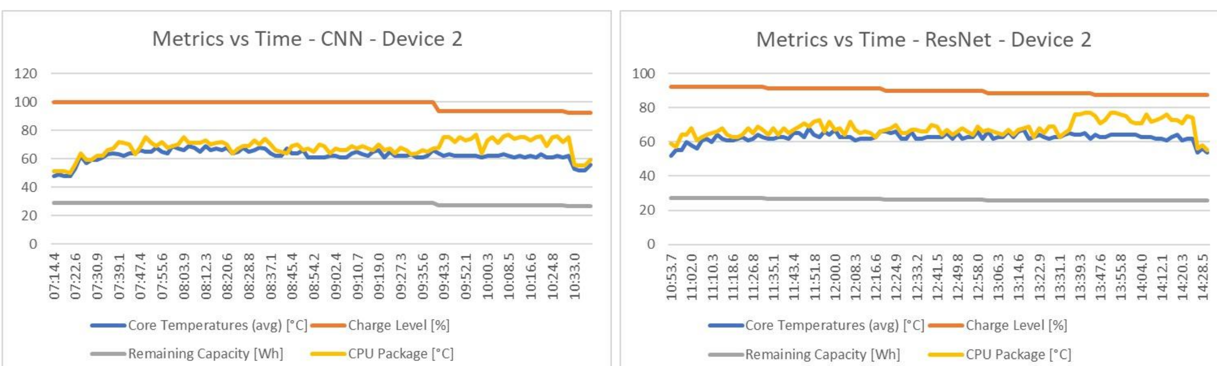

Figure 3: Visualising all Metrics against Time for Device 2

DEVICE 3

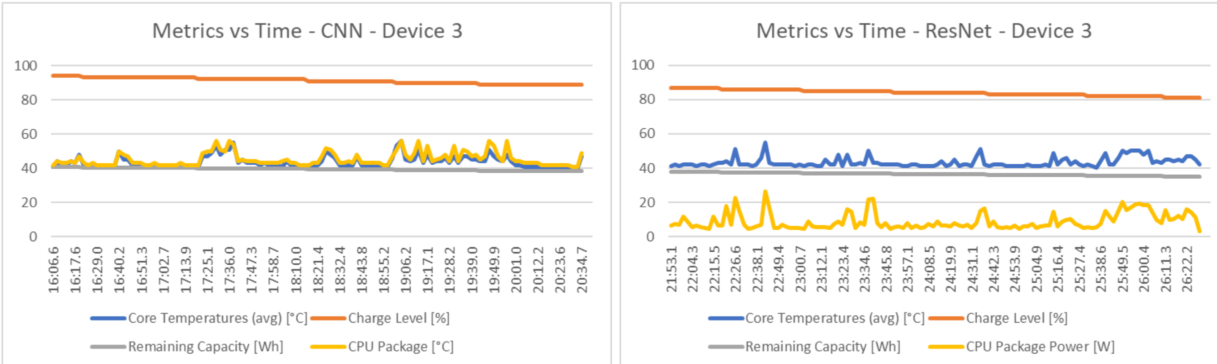

Figure 3: Visualising all Metrics against Time for Device 3

DEVICE 4

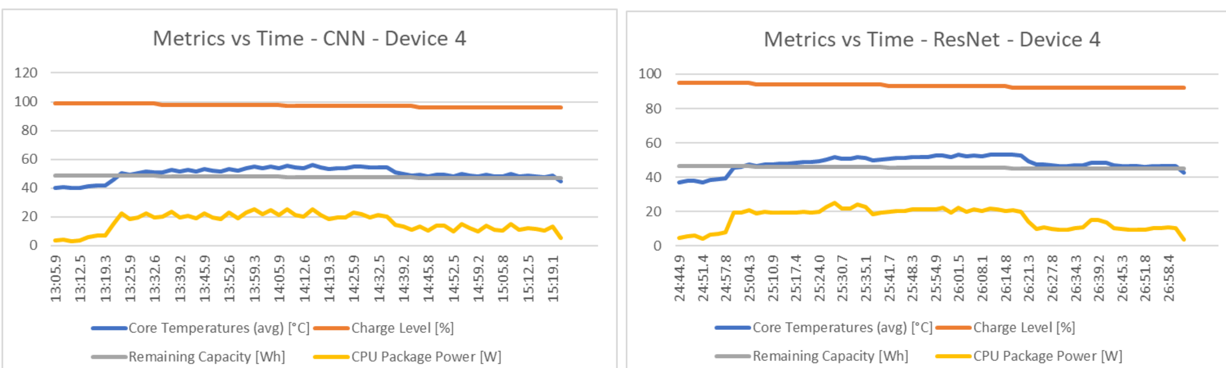

Figure 4: Visualising all Metrics against Time for Device 4

From the below figures you can see the various graphs plotted against time(seconds). To normalise the data, we took the readings from four devices. To ensure that the readings were accurate, we ensured that the devices were put to lowest brightness, the WiFi and Bluetooth connections were switched off and no other program was running in the devices during the commencement of the test.

DEVICE 1

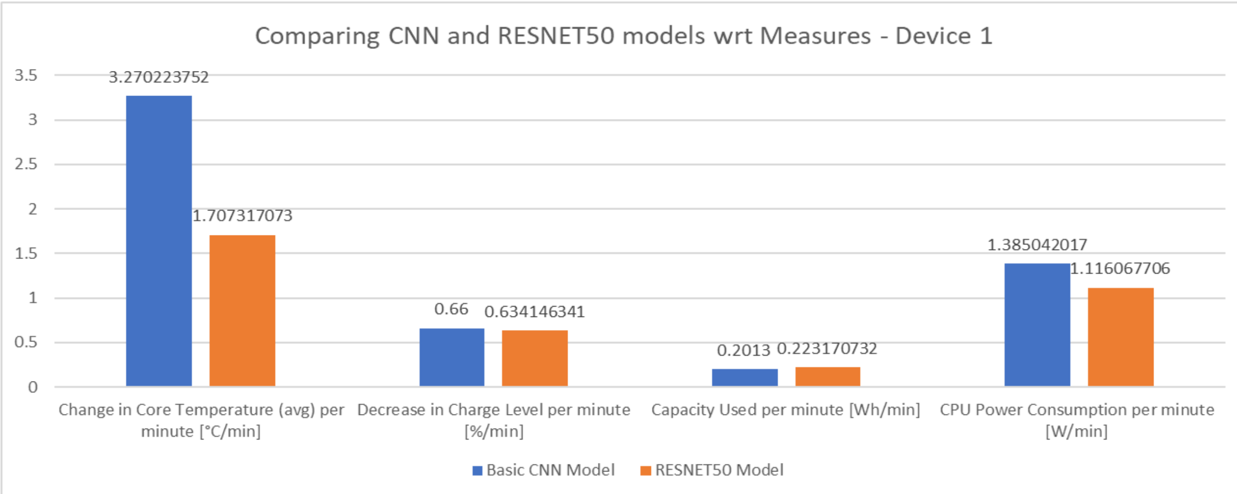

Figure 5: Visualising Calculated Metrics while testing CNN & RESNET50 using Device 1

DEVICE 2

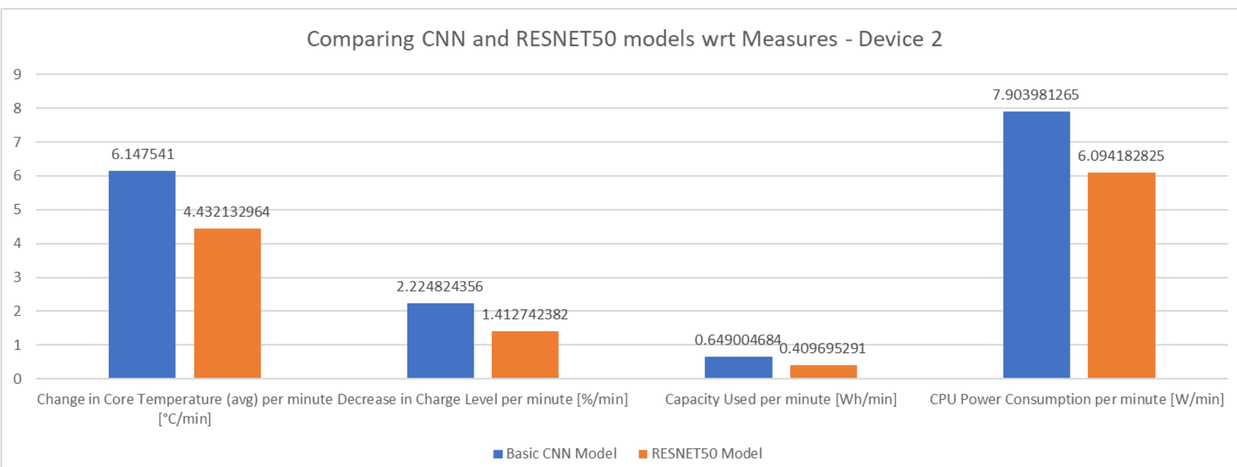

Figure 6: Visualising Calculated Metrics while testing CNN & RESNET50 using Device 2

DEVICE 3

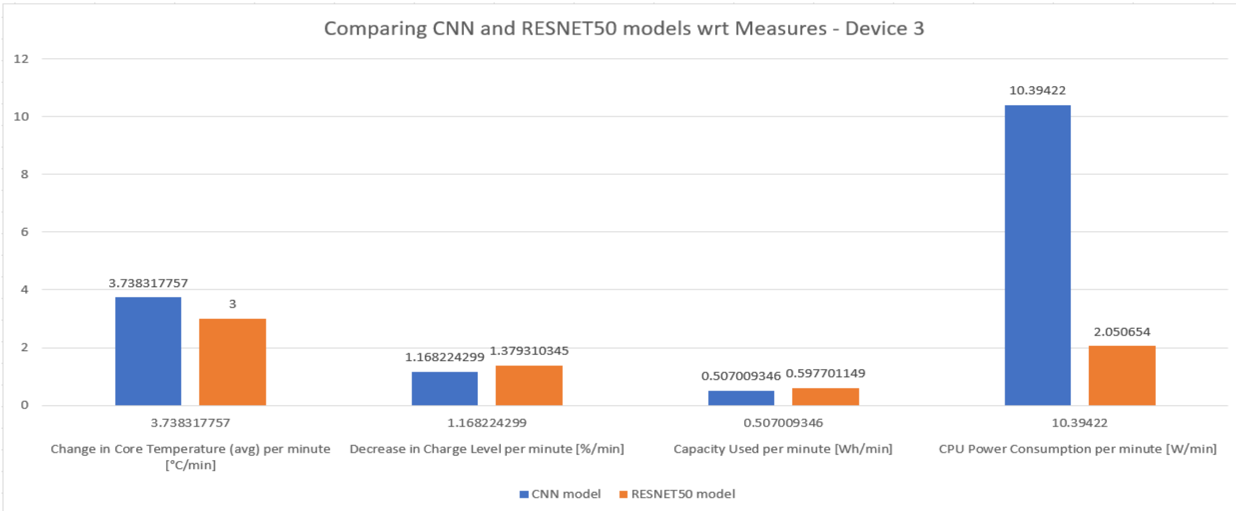

Figure 7: Visualising Calculated Metrics while testing CNN & RESNET50 using Device 3

DEVICE 4

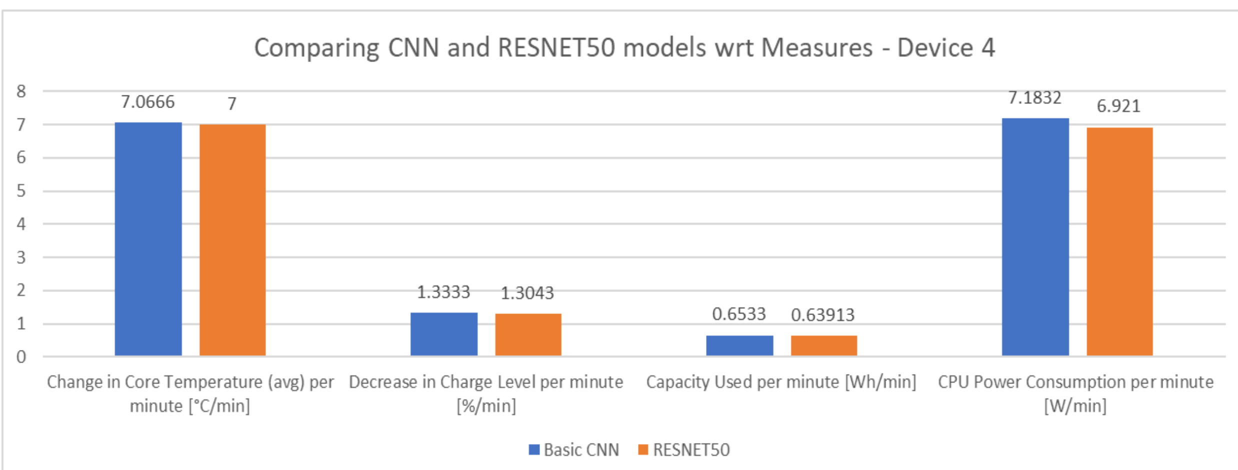

Figure 8: Visualising Calculated Metrics while testing CNN & RESNET50 using Device 4

Device	Model	Change in Core Temperature (avg) per minute [°C/min]	Decrease in Charge Level per minute [%/min]	Capacity Used per minute [Wh/min]	CPU Power Consumption per minute [W/min]
Device 1	Basic CNN Model	3.2702	0.66	0.2013	1.3850
	RESNET50 Model	1.7073	0.6341	0.2231	1.1160
Device 2	Basic CNN Model	6.1475	2.2248	0.6490	7.9039
	RESNET50 Model	4.4321	1.4128	0.4097	6.0941
Device 3	Basic CNN Model	3.7383	1.1682	0.5070	10.3942
	RESNET50 Model	3	1.37931	0.5977	2.0506
Device 4	Basic CNN	7.0666	1.3333	0.6533	7.1832
	RESNET50	7	1.3043	0.6391	6.9210

Table 2: Calculated Metrics values for each Device and Model

From the above bar graph charts, we conclude the following:

1. The change in average core temp per minute is more while executing the basic CNN model than while executing the RESNET50 model. Different devices have different numbers of cores, yet this trend can be observed throughout.
2. The charge level and capacity used per minute across all devices for both the models seems to be similar in value, so both models drain the same amount of battery.
3. CPU power consumption per minute for devices 2 and 3 while testing the basic CNN model is much more than that while testing the RESNET50 model.

Additionally, while testing the two models, we found that the basic CNN model gives more accurate predictions than the RESNET50 model.

From this, we can infer that although the CNN model consumes more resources, the accuracy is much better than RESNET50, making it ideal for emotion detection in iot based systems.

Conclusion

From our findings on this proposed method we concluded that a performance measurement as proposed in this report can be used for real world scenarios where deep learning models are trained for IoT applications. Although the testing method is slightly flawed, we are monitoring the performance metrics at idle power consumption to reduce the impact of system applications

Even though the testing method is slightly flawed, we are monitoring the idle power consumption so as to deduce the difference under load and idle. This will allow us to normalise our results across various platforms and devices. We can use these numbers to estimate battery life on various IOT devices with just their battery capacity. Also the hardware manufactured by large scale industries for IOT devices consist of targeted hardware with driver code very specific to that application. But the laptops we're using to test the algorithms are meant to serve a large variety of users making it the absolute worst case power consumption.

References

1. Mahadik, Ankita & Milgir, Shambhavi & Patel, Janvi & Kavathekar, Vaishali & Jagan, Vijaya Bharathi. (2021). Mood based music recommendation system.
2. J. Tan et al., "Face Detection and Verification Using Lensless Cameras," in IEEE Transactions on Computational Imaging, vol. 5, no. 2, pp. 180-194, June 2019, doi: 10.1109/TCI.2018.2889933.
3. Wen, X. Using deep learning approach and IoT architecture to build the intelligent music recommendation system. Soft Comput 25, 3087–3096 (2021). https://doi.org/10.1007/s00500-020-05364-y
4. H. Kim, J. Ben-othman, L. Mokdad and P. Bellavista, "A Virtual Emotion Detection Architecture With Two-Way Enabled Delay Bound toward Evolutional Emotion-Based IoT Services," in IEEE Transactions on Mobile Computing, vol. 21, no. 4, pp. 1172-1181, 1 April 2022, doi: 10.1109/TMC.2020.3024059.
5. K. Kapoor, K. Gupta, N. Rakesh, R. Gusain, M. Kaur and P. Nand, "IoT based Real-time Face Detection and Recognition System," 2022 9th International Conference on Computing for Sustainable Global Development (INDIACom), 2022, pp. 104-107, doi: 10.23919/INDIACom54597.2022.9763252.
6. K. Choi and G. E. Sobelman, "Optimized Face Detection and Alignment for Low-Cost and Low-Power IoT Systems," 2020 IEEE International Conference on Internet of Things and Intelligence System (IoTaIS), 2021, pp. 129-135, doi: 10.1109/IoTaIS50849.2021.9359713.
7. C. Le and T. K. Mohd, "Facial Detection in Low Light Environments Using OpenCV," 2022 IEEE World AI IoT Congress (AIIoT), 2022, pp. 624-628,
8. doi: 10.1109/AIIoT54504.2022.9817372.
9. Swastik Kumar Sahu, Ram Narayan Yadav, Key facial points recognition using ResNet, Materials Today: Proceedings, Volume 66, Part 8, 2022, Pages 3651-3656, ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2022.07.342.
10. S. Gilda, H. Zafar, C. Soni and K. Waghurdekar, "Smart music player integrating facial emotion recognition and music mood recommendation," 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 2017, pp. 154-158, doi: 10.1109/WiSPNET.2017.8299738.

11. Samuvel, D. J., Perumal, B., & Elangovan, M. (2020). Music recommendation system based on facial emotion recognition. 3C Tecnología. Glosas de innovación aplicadas a la pyme. Edición Especial, Marzo 2020, 261-271. http://doi.org/10.17993/3ctecno.2020.specialissue4.261-271

12. https://ijarcce.com/papers/facial-expression-based-music-recommendation-system/

13. Athavle, Madhuri. (2021). Music Recommendation Based on Face Emotion Recognition. Journal of Informatics Electrical and Electronics Engineering (JIEEE). 2. 1-11. 10.54060/JIEEE/002.02.018.

14. T. Gorasiya, A. Gore, D. Ingale and M. Trivedi, "Music Recommendation based on Facial Expression using Deep Learning," 2022 7th International Conference on Communication and Electronics Systems (ICCES), 2022, pp. 1159-1165, doi: 10.1109/ICCES54183.2022.9835929.

15. Mohini, A. Singh, and A. Khan, "Song recommendation system using facial expression." [Online]. Available: https://ijariie.com/AdminUploadPdf/SONG_RECOMMENDATION_SYSTEM_USING_FACIAL_EXPRESSION_ijariie16138.pdf. [Accessed: 14-Nov-2022].

16. M. Awais et al., "LSTM-Based Emotion Detection Using Physiological Signals: IoT Framework for Healthcare and Distance Learning in COVID-19," in IEEE Internet of Things Journal, vol. 8, no. 23, pp. 16863-16871, 1 Dec.1, 2021, doi: 10.1109/JIOT.2020.3044031.

17. Parthiban, L. and Samy, S.S. (2021). Emotion Detection in IoT-Based E-Learning Using Convolution Neural Network. In Fuzzy Intelligent Systems (eds E. Chandrasekaran, R. Anandan, G. Suseendran, S. Balamurugan and H. Hachimi). https://doi.org/10.1002/9781119763437.ch2

18. P. Chhikara, P. Singh, R. Tekchandani, N. Kumar and M. Guizani, "Federated Learning Meets Human Emotions: A Decentralized Framework for Human–Computer Interaction for IoT Applications," in IEEE Internet of Things Journal, vol. 8, no. 8, pp. 6949-6962, 15 April15, 2021, doi: 10.1109/JIOT.2020.3037207.

19. K. YOKOO, M. ATSUMI, K. TANAKA, H. WANG and L. MENG, "Deep Learning based Emotion Recognition IoT System," 2020 International Conference on

Advanced Mechatronic Systems (ICAMechS), 2020, pp. 203-207, doi: 10.1109/ICAMechS49982.2020.9310135.

20. Richard Jiang, Anthony T.S. Ho, Ismahane Cheheb, Noor Al-Maadeed, Somaya Al-Maadeed, Ahmed Bouridane, Emotion recognition from scrambled facial images via many graph embedding,Pattern Recognition,Volume 67,2017,Pages 245-251,ISSN 0031-3203, https://doi.org/10.1016/j.patcog.2017.02.003.

21. Deepak Kumar Jain, Pourya Shamsolmoali, Paramjit Sehdev,Extended deep neural network for facial emotion recognition,Pattern Recognition Letters,Volume 120,2019,Pages 69-74,ISSN 0167-8655,https://doi.org/10.1016/j.patrec.2019.01.008.

22. "Convolutional Neural Network," DeepAI, 17-May-2019. [Online]. Available: https://deepai.org/machine-learning-glossary-and-terms/convolutional-neural-network. [Accessed: 14-Nov-2022].

23. G. Boesch, "Deep residual networks (ResNet, RESNET50) - 2022 guide," viso.ai, 22-Aug-2022. [Online]. Available: https://viso.ai/deep-learning/resnet-residual-neural-network/. [Accessed: 14-Nov-2022].

24. "ResNet-50: The basics and a quick tutorial," Datagen, 25-Oct-2022. [Online]. Available: https://datagen.tech/guides/computer-vision/resnet-50/. [Accessed: 14-Nov-2022].