Literature review and Model Selection for Binary image classification

I. MODEL SELECTED

Based on the literature reviewed in the 14 papers, the most commonly evaluated and highest-performing model for classification tasks is the ResNet model where ResNet-50 is mentioned in multiple papers as achieving state-of-the-art results.

Several papers achieved their best results using ResNet or variations of it. ResNet was shown to outperform or perform comparably to other popular models like VGGNet, Inception, DenseNet, and MobileNet on various image datasets, also ResNet strikes a balance between high accuracy and computational efficiency. For example, ResNet101 achieves similar accuracy as VGG16 while having 7x fewer parameters. Our key findings that support choosing ResNet:

- It achieved the highest accuracy of 99.3% on the MNIST dataset compared to other models like GoogleNet and VGG16.
- On the PatchCamelyon dataset, ResNet-50 outperformed VGG19 and VGG16 models with 1.0 and 1.2 higher accuracy respectively. It also had a higher AUC-ROC score.
- ResNet-50 attained an impressive 95 accuracy in classifying cancerous vs. non-cancerous lymph node scans.

ResNet-50 is a popular choice due to its effectiveness in training deep networks while addressing the vanishing gradient problem. It utilizes residual blocks that bypass connections to address degradation, improving training efficiency. For classification tasks, ResNet-50 achieves state-of-the-art results on popular image datasets while being easier to optimize than deeper ResNets. The best approach would be to use transfer learning by fine-tuning the pre-trained ResNet-50 model on the target classification dataset [1].

II. LITERATURE REVIEW

Shaveta Arora et al. [2] proposed a research study to apply deep learning techniques to classify brain tumours from MRI scans. They used the BRATS2015 dataset, a collection of MRI brain images, and employed feature extraction through the gray-level co-occurrence matrix (GLCM). They rigorously evaluated several models, including an SVM classifier, random forest classifier, VGG16, Inception_V3, and ResNet. The VGG16 model significantly outperformed traditional machine learning methods, achieving an accuracy of 90.54% on the test set as illustrated in the table I. This suggests that deep learning has the potential to be a powerful tool for brain tumour classification from MRI images.

Model	Train Accuracy	Validation Accuracy	Test Accuracy
SVM Classifier	71.34%	52.56%	50.51%
Random Forest	72.78%	64.3%	64.23%
VGG16	96.3%	92.23%	90.54%
Inception_V3	93.4%	64.8%	63.94%
ResNet	99.7%	82.12%	81.92%

TABLE I: Models Comparison

M. Wang and X. Gong [3] focus on detecting metastatic cancer through image classification, with a specific emphasis on achieving high performance for small patch-level images. The authors propose a novel method based on the ResNet model and evaluate it on the PatchCamelyon (PCam) benchmark dataset. This dataset contains 220,025 samples, consisting of 89,117 cancerous and 130,908 non- cancerous or normal images. The primary model used is the ResNet, known for its effectiveness in training deep networks and addressing gradient-related challenges. Comparative results demonstrate that the ResNet50 model outperforms other models (VGG19 and VGG16) with a 1.0% higher accuracy than VGG19, as well as a 1.2% higher AUC-ROC score and 1.5% higher accuracy than VGG16 as illustrated in the table II . Evaluation metrics encompass the AUC-ROC score, assessing the classifier's ability to differentiate between cancerous and non-cancerous classes, and accuracy, providing an overall measure of the model's correctness in image classification.

Models	AUC (ROC Score)	Accuracy
Vgg16	0.951	0.957
Vgg19	0.955	0.962
ResNet50	0.963	0.972

TABLE II: Models Comparison

Jun Hur and Haewoon Nam's [4] study focused on distinguishing between human and animal targets using radar signals from the MAFAT Radar Challenge dataset, which comprises signals recorded by ground Doppler pulse radar. These signals were segmented into 32 time-unit segments, each consisting of a 32x128 matrix, with each segment representing either a human or animal moving within the radar's range. The study employed three deep learning models for classification: CNN, U-Net, and Res-UNet. Their performance was evaluated based on their ability to correctly identify humans and animals. The Res-UNet model outperformed the others, underscoring its potential for enhancing radar target detection as illustrated in the table III.

Model Name	Human Accuracy	Animal Accuracy
CNN	95.73%	97.42%
U-Net	98.26%	95.25%
Res-UNet	98.27%	97.71%

TABLE III: Models Comparison

K. Dong [5] aimed to perform image classification using MobileNetV2 and compare its performance with MobileNetV1. The dataset used was the large-scale ImageNet dataset consisting of over 1 million images belonging to 1000 classes. Results showed that MobileNetV2 achieved 72.0% top-1 accuracy on ImageNet with 301M multiply-adds, an improvement over MobileNetV1 which attained 70.6% accuracy with 569M multiply-adds. When compared to other models, MobileNetV2 achieved similar accuracy as ResNet-101 but with 5x fewer computations, and approached VGG-16 accuracy with 23x fewer multiply-adds, demonstrating its improved efficiency as illustrated in the table IV.

Model	Top-1 Accuracy	Multiply-Adds (M)
MobileNetV1	70.6%	569
MobileNetV2	72.0%	301
ResNet-101	75.2%	1550
VGG-16	74.5%	15300

TABLE IV: Models Comparison

S. Sharma [6] review of various deep learning models for image classification including AlexNet, ResNet, GoogleNet and VGG16/19. The models are evaluated on several popular databases like MNIST, ImageNet, fruit and leaf classification datasets and hyperspectral image datasets. The goal of the study is to compare the performance of these models in terms of different metrics like accuracy, sensitivity and specificity. The results show that ResNet achieved the highest accuracy of 99.3% on MNIST dataset while GoogleNet performed best for vegetable and fruit classification with 94.5% accuracy. VGG16 achieved an accuracy of 98.4% for leaf disease classification as illustrated in the table V.

Model name	Accuracy
ResNet	99.3%
GoogleNet	94.5%
VGG16	98.4%

TABLE V: Models Comparison

C. Ma [7] researched image classification methods based on deep convolutional neural networks (DCNN). It used the CIFAR-10 image dataset to train and evaluate different DCNN models including VGG-16, Inception v3, ResNet and DenseNet. The classification accuracy was used as the evaluation metric. VGG-16 achieved an accuracy of 93.4%, Inception v3 achieved 94.2%, ResNet achieved 95.1% and DenseNet achieved the highest accuracy of 95.8% as illustrated in the table VI. The results were compared between these DCNN models to analyze their performance on image classification tasks.

Model name	Accuracy
VGG-16	93.4%
Inception v3	94.2%
ResNet	95.1%
DenseNet	95.8%

TABLE VI: Models Comparison

Liang Zhang et al. [8] proposed a deep learning-based approach to classify rocket images derived from aerospace optical equipment. They developed a binary classification model using the ResNet18 framework and modified the binary cross entropy loss function to improve generalization performance on difficult images. The model was trained on a dataset of 2,689 images and achieved 100% precision, 83.33% recall, and 95.87% F1 score on the test set of 97 images, which included 24 rocket images as illustrated in the table VII. Four rocket images were misclassified as non-rocket images. The model's high performance demonstrates the potential of deep learning for rocket image classification.

Model	Precision	Recall	F1 score
ResNet18	100%	83.33%	95.87%

TABLE VII: Model Results

I. J. Jin [9] presents an innovative diagnostic system utilizing deep learning and IR thermography to classify system conditions, offering real-time detection of abnormalities and accidents. Thermal images captured during thermal-hydraulic tests in nuclear power plants were used to train a Convolutional Neural Network (CNN) for component diagnosis, system-wide assessment, and accident classification. The choice of CNN architecture, including AlexNet, GoogleNet, VGGNet, ResNet, and DenseNet, played a crucial role in determining classification time efficiency, assessed through Multiply and Accumulate per second (MACs) shown comparison in the table VIII. The study optimized hyperparameters with the Adaptive Momentum Estimation (Adam) algorithm, a learning rate of 0.001, a batch size of 4, and 100 training epochs without pre-trained weights, using PyTorch as the platform for deep learning application. This approach enables rapid and precise condition monitoring, valuable for enhancing safety in nuclear power plants.

H. Tang [10] provides a review and qualitative comparison of key CNN models and modules for image classification. Several seminal models including LeNet, AlexNet, Inception, VGGNet and ResNet are discussed. The models are compared based on classification accuracy and computational efficiency. Specific datasets are not mentioned, rather a general discussion is provided on model architectures, components like pooling, activation

Architectures	Parameters	MACs
ANN (200, 200, 200)	60,373,204.00	30,186,000.00
ANN (300, 300, 300)	90,679,804.00	45,339,000.00
ANN (400, 400, 400)	121,066,404.00	60,532,000.00
ANN (500, 500, 500)	151,533,004.00	75,765,000.00
AlexNet	124,756,688.00	1,135,256,096.00
VGGNet-16	268,554,372.00	15,466,176,512.00
VGGNet-19	279,178,884.00	19,627,974,656.00
GoogLeNet	20,640,428.00	1,575,133,184.00
ResNet-18	22,355,076.00	1,813,562,368.00
ResNet-34	42,571,396.00	3,663,250,432.00
ResNet-50	47,024,260.00	4,087,140,352.00
ResNet-101	85,008,516.00	7,799,361,536.00
ResNet-152	116,295,812.00	11,511,582,720.00
DenseNet-101	1,580,564.00	416,052,624.00

TABLE VIII: Models Architectures

functions and dropout, and performance trade-offs. Key metrics examined are top-1 classification accuracy and multiply-add operations for efficiency. There are no results reported for any models along with the metrics used. The paper lacks critical experimental details needed for a meaningful evaluation of the methods presented.

V. Tiwari [11] classifies images into categories like living/non-living and further into more granular classes using deep neural networks. The authors train and evaluate models on a dataset of over 45,000 images across 5 classes including nature images from VUB, animal images from Kaggle, group photos from Cornell, selfies from UCF, and vehicles from Stanford. The main model proposed is a modified 16-layer VGG architecture with a custom classifier. They also experiment with a baseline CNN, 3-block VGG, and data augmentation. The best accuracy of 98.97% is achieved by the VGG-16 model with softmax activation and categorical cross-entropy loss. Other metrics reported are training time and model size as illustrated in the table IX.

Model	Accuracy	Loss Function
Baseline CNN	55.075%	binary cross-entropy
VGG 3	74.561%	binary cross-entropy
VGG 3 + Data Augmentation	61.404%	binary cross-entropy
VGG 16	98.97%	categorical cross-entropy

TABLE IX: Models Comparison

A. S. A. Nisha, [12] proposes a deep learning model for medical image classification which incorporates Convolutional Neural Network (CNN), Naive Bayes, Support Vector Machine (SVM) and

Multilayer Perceptron (MLP). The datasets used for the evaluation of the model are HIS2828 and ISIC2017. Various classification algorithms such as CNN, Naive Bayes, SVM and MLP are used to classify the images and results are compared with current techniques. A classification accuracy of 97% and 96% is achieved respectively using the proposed model which is higher than existing techniques, thus proving the effectiveness of the proposed approach as illustrated in the table X.

Model name	Accuracy
CNN	97.5%
Naive Bayes	97.3%
MLP	98.92%
SVM	97.79%

TABLE X: Models Comparison

Reham S. Saeed and Bushra K. Oleiwi [13] aimed to develop a deep learning model for the binary classification of COVID-19 based on Chest X-ray images. The dataset used was obtained from public sources on Kaggle and consisted of positive COVID-19 cases and normal cases. The researchers utilized a Convolutional Neural Network (CNN) model for the classification task. The results showed promising performance, with an accuracy of 96.68%, recall of 94.12%, F1 Score of 93.8%, specificity of 97.61%, and precision of 93%. The performance of the model was evaluated using these metrics as illustrated in the table XI.

Model	**Recall**	**Accuracy**	**Precision**	**F1 Score**
CNN	94.12%	96.68%	93.49%	93.80%

TABLE XI: Performance Metrics

K. Gupta and V. Bajaj [14] aim to automate COVID-19 screening via deep learning on CT-scan images, enhancing early diagnosis and containment. The method includes resizing images, transfer learning, and a lightweight DLM. In biomedical signal/image analysis, the absence of standardized DLMs prompts careful model selection based on performance and efficiency. DarkNet19, a pre-trained model for real-time object detection, and MobileNetV2, optimized for mobile devices, are utilized. The evaluation involves eight parameters, highlighting DarkNet19's highest average classification accuracy of 98.91%. DarkNet19 achieves this using the least training time (238 minutes) compared to MobileNetV2 (177 minutes) and the proposed DLM (94 minutes) as illustrated in the table XII. The proposed DLM also exhibits minimal loss percentage, and DarkNet19 covers the largest area. Importantly, this system's adaptability extends to screening other diseases.

Parameters	**DLM**	**MobileNetV2**	**DarkNet19**
ACC (%)	95.93	97.62	98.91
SEN (%)	95.05	97.76	98.96
SPE (%)	96.82	97.47	98.86
F1-score	0.96	0.98	0.99
PRC (%)	96.82	97.52	98.88
NPV (%)	95.04	97.71	98.94
FM	0.96	0.98	0.99
AUC (%)	98.94	99.67	99.89

TABLE XII: Performance Comparison

Ghadekar et al. [15] present a system for detecting cancer in histopathological scanned images using deep learning techniques. It utilizes the ResNet 50 model, employing transfer learning in Convolutional Neural Networks. The goal is to classify whether a given lymph node scan is cancerous or non-cancerous. The dataset used is the PatchCamelyon benchmark dataset from Kaggle, containing 277,483 images. The primary model, ResNet 50, is known for its effectiveness in image classification tasks. The system achieved an impressive 95% accuracy in classifying cancerous and non-cancerous lymph node scans, with no specific details on other comparisons provided. The main evaluation metric used is accuracy, measuring the proportion of correctly classified instances in of the dataset.

References

[1] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

[2] S. Arora and M. Sharma, "Deep Learning for Brain Tumor Classification from MRI Images," 2021 Sixth International Conference on Image Information Processing (ICIIP), Shimla, India, 2021, pp. 409-412, doi: 10.1109/ICIIP53038.2021.9702609.

[3] M. Wang and X. Gong, "Metastatic Cancer Image Binary Classification Based on Resnet Model," 2020 IEEE 20th International Conference on Communication Technology (ICCT), Nanning, China, 2020, pp. 1356 - 1359, doi: 10.1109/ICCT50939.2020.9295797.

[4] J. Hur and H. Nam, "Performance Comparison of Moving Target Classification based on Deep Learning," 2022 13th International Conference on Information and Communication Tech-

nology Convergence (ICTC), Jeju Island, Korea, Republic of, 2022, pp. 1533-1535, doi: 10.1109/ICTC55196.2022.9952676.
[5] K. Dong, C. Zhou, Y. Ruan and Y. Li, "MobileNetV2 Model for Image Classification," 2020 2nd International Conference on Information Technology and Computer Application (ITCA), Guangzhou, China, 2020, pp. 476-480, doi: 10.1109/ITCA52113.2020.00106.
[6] S. Sharma and K. Guleria, "Deep Learning Models for Image Classification: Comparison and Applications," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2022, pp. 1733-1738, doi: 10.1109/ICACITE53722.2022.9823516.
[7] C. Ma, S. Xu, X. Yi, L. Li and C. Yu, "Research on Image Classification Method Based on DCNN," 2020 International Conference on Computer Engineering and Application (ICCEA), Guangzhou, China, 2020, pp. 873-876, doi: 10.1109/ICCEA50009.2020.00192.
[8] L. Zhang, Z. Chen, J. Wang and Z. Huang, "Rocket Image Classification Based on Deep Convolutional Neural Network," 2018 10th International Conference on Communications, Circuits and Systems (ICCCAS), Chengdu, China, 2018, pp. 383-386, doi: 10.1109/ICCCAS.2018.8769176.
[9] I. J. Jin, D. Y. Lim, and I. C. Bang, "Deep-learning-based system-scale diagnosis of a nuclear power plant with multiple infrared cameras," Nuclear Engineering and Technology, vol. 55, no. 2, pp. 493-505, 2023. doi: https://doi.org/10.1016/j.net.2022.10.012.
[10] H. Tang, "Image Classification based on CNN: Models and Modules," 2022 International Conference on Big Data, Information and Computer Network (BDICN), Sanya, China, 2022, pp. 693-696, doi: 10.1109/BDICN55575.2022.00134.
[11] V. Tiwari, C. Pandey, A. Dwivedi and V. Yadav, "Image Classification Using Deep Neural Network," 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Greater Noida, India, 2020, pp. 730-733, doi: 10.1109/ICACCCN51052.2020.9362804.
[12] A. S. A. Nisha, V. Enireddy, T. Bernatin, K. C, D. V. Babu and N. A. Vignesh, "Deep Learning based Medical Image Classification," 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2022, pp. 974-978, doi: 10.1109/ICSSIT53264.2022.9716383.
[13] R. S. Saeed and B. K. Oleiwi, "Deep learning model for binary classification of COVID-19 based on Chest X-Ray," 2023 15th International Conference on Developments in eSystems Engineering (DeSE), Baghdad & Anbar, Iraq, 2023, pp. 45-49, doi: 10.1109/DeSE58274.2023.10099555.
[14] K. Gupta and V. Bajaj, "Deep learning models-based CT-scan image classification for automated screening of COVID-19," Biomedical Signal Processing and Control, vol. 80, p. 104268, 2023. doi: https://doi.org/10.1016/j.bspc.2022.104268.
[15] P. Ghadekar, A. Khandelwal, P. Roy, A. Gawas and C. Joshi, "Histopathological Cancer Detection using Deep Learning," 2021 International Conference on Artificial Intelligence and Machine Vision (AIMV), Gandhinagar, India, 2021, pp. 1 - 6, doi: 10.1109/AIMV533 13.2021.9670 991.