Image Classification Challenge with Deep Learning Models

I. Machine Learning

Machine learning has grown to be a critical factor in today's real-world applications, changing the way many areas like health, banking and transportation operate. Recent advancements in Large Language Models (LLMs) have expanded applications to include intelligent chatbots. This report deals with the prospective applications of machine learning (ML), show the differences between different ML approaches and discuss the existing barriers and their solutions.

A. *Applications of Machine Learning*

Machine learning is remaking healthcare by allowing researchers to dig into big piles of patient data and establish relationships and correlations showing up diseases and conditions. Through high level algorithms and statistical methods medical specialists can make an accurate diagnoses, form a tailored treatment course and improve patient's results.

One of the main domains of AI utilization is discriminating diseases. Through the usage of machine learning to process images and electronic health records [1], the algorithm can recognize sublime patterns and not so evident irregularities which ultimately tell a specific disease. For instance, a machine learning technique was employed to create computer-aided detection systems with high levels of precision for breast and lung cancer.

In addition, machine learning contributes much to both the drug discovery and development [2]. Analysing datasets of chemical compounds, genes, and clinical trial results, algorithms can pick out drugs made from compounds that were previously used for other diseases, or employ existing drugs for new diseases, thus cutting the time and costs.

Machine learning also shows superiority in customized treatment planning or decision-making. Through analysing the historical medical records and genomic information, algorithms examine the most appropriate treatment techniques and probability of an outcome. Such process will guide them in providing more clarified answers and better improve patient outcomes.

Machine learning has a great deal of potential; however, issues such as: access to high-quality data, algorithm bias, and regulation need to be resolved to be able to fully realize machine learning's advantage in healthcare.

Moreover, ML has begun to make a significant impact on banking as it increasingly replaces the traditional methods for fraud detection [3] and credit risk assessment. ML algorithms allow the financial authorities to quickly scan through big amounts of transaction data in search of any suspicious activities or fraud which is happening right in a moment. ML models analyse transaction patterns for unexpended intensity, mysterious geo-locations and erratic cash out. In ML systems, historical fraud patterns are constantly being examined and evaluated, and through feedback, they are continue to grow smarter and stay a step ahead of the ever-evolving cybercriminal strategies. ML algorithms not only helps to assess risk and manage portfolio through the analysis of market trends and economic indications but also can be a source of new investment ideas. This reduces risks giving banks room to land sound investment decisions. ML enhanced risk assessment models are capable of identifying lending practice vulnerabilities, and thus credit risks will decline and better –screened loan application process will be ensured. On the whole, ML assimilation in banking industry step ups up fraud detection, improves risk assessment, and simplifies the processes of decision making in order to gain efficiency and profitability.

Machine learning (ML) is taking the whole transportation system to a new level through route optimization, scheduling improvement, improvement in transportation safety, and efficient transportation system. ML technology is the base of the development of the block of autonomous vehicles. The algorithms of this system are sophisticated and capable to make quick decisions in real world.

The sensor systems of autonomous vehicles consist of cameras, LiDAR, radar, and GPS devices and receive all sorts of information about their surroundings. These algorithms process the data to give computers a perception concerning what is around them and the ability to identify objects and understand them. This feature in self-driving vehicles keeps them proactive and allows them to rapidly respond to emergencies and road chaotic situations.

Furthermore, through ML, route is conveniently planned and logistics are facilitated by studying weather condition, historical traffic patterns [4] and the real time data. Elevated routing and schedule are allocated, achieving to shorten distances travelled and amount of fuel used as well as crowd control and air pollution reduction.

Besides, AI improves safety better than humans as it has an ability of detecting lane departure, collision and tired driving behaviour. In this case, the alarm can be sent to drivers, and some urgent measures can even be taken into consideration to stop the crash.

Algorithms based on ML can do the job of looking at information coming from multiple sources and find out traffic congestion patterns to make roads more traffic smooth and efficient. Through timely adjusting traffic signals and predictive messages to drivers, the congestion may be solved and an outlook is achieved onto urban motion.

In conclusion, machine learning has emerged as a transformational force in a variety of industries, including healthcare, banking, and transportation. Machine learning has prospered and is adapted in various operating sectors, and with a high chance that future of technology-driven solutions will be improved as new inventions are unveiled.

B. *Types of Machine Learning Methods*

Three primary ML methods include supervised learning, unsupervised learning, and reinforcement learning. Supervised learning excels with labeled data but requires costly data acquisition. Unsupervised learning can uncover hidden patterns but may struggle with data complexity.

Reinforcement learning is ideal for decision-making tasks but often requires extensive computational resources [5].

1) *Supervised Learning*

a) *Pros*

This method is suitable for tasks with labelled data, such as classification and regression and can generalize well to unseen data if the model is properly trained.

b) *Cons*

This method requires a substantial amount of labelled data for training and has limited applicability to problems without labelled data.

2) *Unsupervised Learning*

a) *Pros*

This method can identify hidden patterns and structures in unlabelled data and it can handle high-dimensional data effectively.

b) *Cons*

Results of this method may be subjective and difficult to interpret and it has limited capability to generalize to new data compared to supervised learning.

3) *Reinforcement Learning*

a) *Pros*

This method can learn optimal strategies through trial and error interactions with the environment and it is well-suited for complex tasks such as game playing and robotics.

b) *Cons*

This method requires extensive exploration and experimentation to learn optimal policies and it is prone to instability and slow convergence, especially in high-dimensional or continuous action spaces.

C. *Challenges and Potential Solutions*

The main problems with machine learning nowadays are related to data privacy, the model interpretation, and the need for large datasets.

Privacy of data is a great concern because machine learning systems usually require a vast amount of data to train the model, during which the sensitive information might be at risk. In order to resolve this, new algorithms, such as Federated Learning [6] and Differential Privacy [7], are being designed. The function of Federated Learning is to train models across many randomized devices, Differential Privacy on the other hand makes the data less identifiable and brings in noise to it.

The absence of model interpretability is another challenge as many models are being viewed black boxes rather than being clear models. The approach is XAI [8] i.e. .Explainable AI which intends to create model that give easily understandable explanation of its resolutions.

Eventually, the most persistent challenge is that of the large datasets. Such techniques like data augmentation [9] that involves creating of new data by applying transformations to existing dataset or transfer learning [10] that is based on pre-trained models for new tasks are being used to deal with this problem. These techniques can dramatically reduce the cost of training by saving data space, thus making machine learning more affordable and practical.

II. MACHINE LEARNING APPLICATION

A. *Introduction*

Image classification is one of the most important task in machine learning and for this task, deep learning models are used to classify provided images. These images are classified into 10 different classes. Pre-trained deep learning models are used as base models and image dataset is used to finetune these models to accurately classify images.

B. *Data and preliminary analysis*

The dataset consists of 2 subdirectories, namely "train" and "val". Each of these directories contains 10 subdirectories, representing classes. Within these directories, there are a total of 9,469 training images and 3,925 validation images stored.

During initial data analysis, dimensions of each images were plot to analyse variation in dimensions of data and it was analysed that there is significant variation in image sizes but most of the images are under 1000x1000 pixels. Figure 1 shows the variation in image sizes and Figure 2 shows distribution of dataset image sizes.

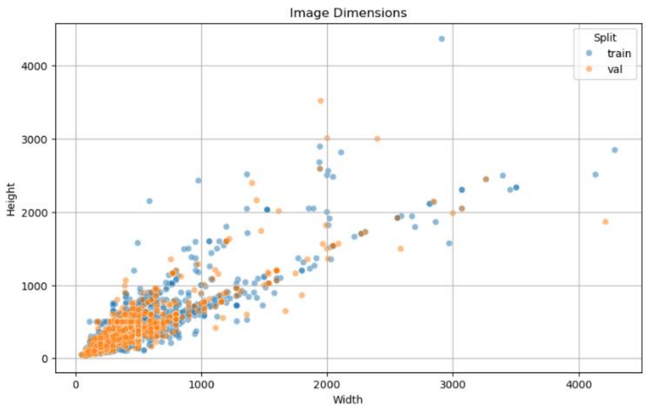

Figure 1 Image height and width scatter plot

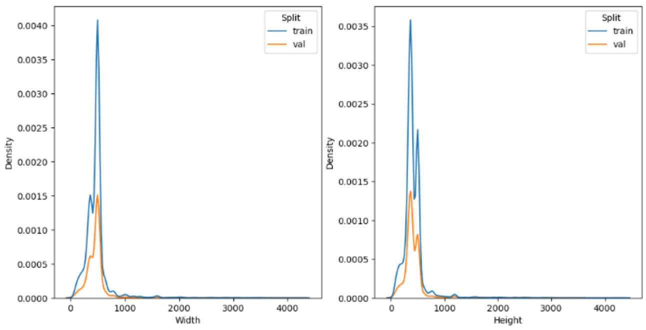

Figure 2 Image sizes distribution

Both plot figures shows most of the images are under the size of 1000x1000 pixels.

Due to such variation in dataset, images are resize to 224x224 pixel for 2 reasons. Firstly, most of the models which we used require images in this dimension and second is to make every image to same size for consistency in dataset.

C. *Methods*

For image classification task, 11 different pretrained models were used. These models are already trained on large dataset and they demonstrated excellent performance. This method also reduces the computational cost as models trained from scratch are computationally inefficient and requires substantial time and resources (CPU, GPU).

For this assignment, Tensorflow [11] is chosen as AI framework due to the reason, its performance is often preferred for image classification problems being due to its flexibility, scalability and performance improvements. Its flexible API provides developers with the power to create a variety of neural network architectures that can be adapted to serve different needs ranging from very basic to advanced complexities. TensorFlow provides the possibility of scaled cloning from prototyping for deployment and efficiently handles the training of deep networks on large data volume. TensorFlow with GPU and TPU acceleration support offers better speed optimization during training. Also, the support of ready-to-use pre-trained models and tools are available through the TensorFlow Hub and TensorFlow Model Garden. TensorFlow is integrated well with the popular libraries that makes the workflow smooth while its active community provides enough support and resource needed by the developers. That is why it has become the major go-to technology of the moment for the solution of image classifications.

Following pre-trained models are used in this project,

ResNet50, ResNet101, ResNet152: These models are based on the ResNet architecture [12]. They employ residual connections to enable the training of very deep neural networks.

VGG19, VGG16: The VGG models are based on the VGG architecture [13]. They consist of multiple convolutional layers followed by max-pooling layers.

EfficientNetB0 [14], EfficientNetV2B0, EfficientNetV2L [15]: The EfficientNet models are based on the EfficientNet architecture. They use a compound scaling method to balance model depth, width, and resolution for optimal performance.

DenseNet121, DenseNet169, DenseNet201: These models are based on the DenseNet architecture [16]. They connect each layer to every other layer in a feed-forward fashion, leading to strong feature reuse and compact models.

D. Experiments

Two variations of input data are employed for experimentation purposes. One approach involves passing the complete data without any random data augmentation or preprocessing other than resizing and another approach is to apply data augmentation on images and using regularization method like 50% Dropout. This approach facilitates not only the comparison of experiments among different models but also the assessment of the impact of data augmentation on the results.

Models are trained using accuracy metric for 10 epochs to find best performing model. Table 1 and Table 2 shows the accuracy results of experiment 1 and experiment 2 respectively on training dataset.

Table 1 Training accuracy of Exp. 1 on different models

Models	Accuracy
DenseNet121_accuracy	77.79%
DenseNet169_accuracy	75.83%
DenseNet201_accuracy	78.95%
EfficientNetB0_accuracy	99.94%
EfficientNetV2B0_accuracy	99.98%
EfficientNetV2L_accuracy	99.90%
ResNet101_accuracy	100.00%
ResNet152_accuracy	100.00%
ResNet50_accuracy	99.98%
VGG16_accuracy	99.57%
VGG19_accuracy	99.48%

Table 2 Training accuracy of Exp. 2 on different models

Models	Accuracy
DenseNet121_accuracy	55.29%
DenseNet169_accuracy	52.68%
DenseNet201_accuracy	59.77%
EfficientNetB0_accuracy	94.63%
EfficientNetV2B0_accuracy	96.39%
EfficientNetV2L_accuracy	98.04%
ResNet101_accuracy	89.91%
ResNet152_accuracy	90.43%
ResNet50_accuracy	89.86%
VGG16_accuracy	86.44%
VGG19_accuracy	85.70%

It is evident from the experiments that accuracy values of experiment 1 is better than experiment 2 even perfect values but 100% accuracy shows overfitting of training data i.e. model has learnt training dataset completely. Models trained during experiment 2, with the use of data augmentation and Dropout produces more generalized results. Overall, in experiment 1, ResNet101 and ResNet152 perform the best with an accuracy of 100% and in experiment 2 EfficientNetV2L performs the best with an accuracy of 98.04%. DenseNet performs poorly in both experiments.

After training, validation is passed to the model for validation accuracy. Table 3 and Table 4 shows the validation set accuracies of experiment 1 and experiment 2 respectively.

Table 3 Validation accuracy of Exp 1 on different models

Models	Accuracy
DenseNet121_accuracy	72.76%
DenseNet169_accuracy	71.11%
DenseNet201_accuracy	73.17%
EfficientNetB0_accuracy	99.29%
EfficientNetV2B0_accuracy	99.59%
EfficientNetV2L_accuracy	99.69%
ResNet101_accuracy	97.17%
ResNet152_accuracy	97.68%
ResNet50_accuracy	97.53%
VGG16_accuracy	93.96%
VGG19_accuracy	94.55%

Table 4 Validation accuracy of Exp 2 on different models

Models	Accuracy
DenseNet121_accuracy	35.08%
DenseNet169_accuracy	33.20%

DenseNet201_accuracy	27.69%
EfficientNetB0_accuracy	98.62%
EfficientNetV2B0_accuracy	99.41%
EfficientNetV2L_accuracy	99.64%
ResNet101_accuracy	95.19%
ResNet152_accuracy	94.70%
ResNet50_accuracy	93.89%
VGG16_accuracy	91.87%
VGG19_accuracy	91.39%

Same pattern is observed in validation set accuracy, experiment 1 has better accuracy values than experiment 2. Also, it is noteworthy that validation accuracy in experiment 2 is greater than training accuracy except for DenseNet. This is due to the usage of Dropout during training only and during where certain neurons are deactivated.

E. Reflection

Usage of pre-trained models reduces the time to train model on custom dataset and computationally it is efficient to perform. Additionally, data augmentation create modified images from the dataset which improves model training and prevent overfitting. However, if resources are not constrained, training a model from scratch is also an option worth considering.

III. References

[1] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.

[2] S. Dara, S. Dhamercherla, S. S. Jadav, C. M. Babu, and M. J. Ahsan, "Machine Learning in Drug Discovery: A Review," *Artif. Intell. Rev.*, vol. 55, no. 3, pp. 1947–1999, Mar. 2022, doi: 10.1007/s10462-021-10058-4.

[3] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, "Credit card fraud detection using machine learning techniques: A comparative analysis," in *2017 International Conference on Computing Networking and Informatics (ICCNI)*, Lagos: IEEE, Oct. 2017, pp. 1–9. doi: 10.1109/ICCNI.2017.8123782.

[4] J. Rzeszótko and S. H. Nguyen, "Machine Learning for Traffic Prediction," *Fundam. Informaticae*, vol. 119, no. 3–4, pp. 407–420, 2012, doi: 10.3233/FI-2012-745.

[5] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, Jul. 2015, doi: 10.1126/science.aaa8415.

[6] P. Kairouz *et al.*, "Advances and Open Problems in Federated Learning." arXiv, Mar. 08, 2021. Accessed: Mar. 08, 2024. [Online]. Available: http://arxiv.org/abs/1912.04977

[7] C. Dwork and A. Roth, "The Algorithmic Foundations of Differential Privacy," *Found. Trends® Theor. Comput. Sci.*, vol. 9, no. 3–4, pp. 211–407, 2013, doi: 10.1561/0400000042.

[8] V. Belle and I. Papantonis, "Principles and Practice of Explainable Machine Learning," *Front. Big Data*, vol. 4, p. 688969, Jul. 2021, doi: 10.3389/fdata.2021.688969.

[9] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J. Big Data*, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.

[10] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *J. Big Data*, vol. 3, no. 1, p. 9, Dec. 2016, doi: 10.1186/s40537-016-0043-6.

[11] M. Abadi, "TensorFlow: learning functions at scale," in *Proceedings of the 21st ACM SIGPLAN International Conference on Functional Programming*, Nara Japan: ACM, Sep. 2016, pp. 1–1. doi: 10.1145/2951913.2976746.

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition." arXiv, Dec. 10, 2015. Accessed: Mar. 11, 2024. [Online]. Available: http://arxiv.org/abs/1512.03385

[13] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition." arXiv, Apr. 10, 2015. Accessed: Mar. 11, 2024. [Online]. Available: http://arxiv.org/abs/1409.1556

[14] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks." arXiv, Sep. 11, 2020. Accessed: Mar. 11, 2024. [Online]. Available: http://arxiv.org/abs/1905.11946

[15] M. Tan and Q. V. Le, "EfficientNetV2: Smaller Models and Faster Training." arXiv, Jun. 23, 2021. Accessed: Mar. 11, 2024. [Online]. Available: http://arxiv.org/abs/2104.00298

[16] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks." arXiv, Jan. 28, 2018. Accessed: Mar. 11, 2024. [Online]. Available: http://arxiv.org/abs/1608.06993