RESEARCH ARTICLE

DEEP LEARNING-BASED CLASSIFICATION OF PHARMACEUTICAL STARCHES USING MICROSCOPY AND TRANSFER LEARNING

*Joy I. Odimegwu, **Marvellous Ajala, *Blessing Oko, *David Oba-Fidelis, *Joycelyn Iyasele,
*Department of Pharmacognosy, Faculty of Pharmacy, University of Lagos, Nigeria.
**Magami Open Sciences Initiative.Lagos, Nigeria.
Corresponding author: jodimegwu@unilag.edu.ng

ABSTRACT

The accurate identification of starch sources in drug discovery is critical especially in Pharmaceutical drug formulations, where excipients variability can affect drug stability, bio-availability, and therapeutic efficacy. This study applies Artificial Intelligence (AI), specifically Convolutional Neural Networks (CNNs), to classify microscopical images of starch granules from 9 botanical sources. A dataset of 889 images was curated using smart phones and light microscopes, and two pretrained models; ImageNet and MicroNet which were fine-tuned for classification. The Image Net-based model achieved an accuracy of 81.2%, outperforming MicroNet 61.0% across precision, recall, and Function measure (F1-score) metrics. Miscalculation patterns were observed in morphologically similar starches such as millet and maize, highlighting the need for enhanced imaging protocols.

This AI-assisted approach offers a scalable, objective alternative to manual (human eyes) starch identification, with implications for quality control in pharmaceutical and nutraceutical industries.

Keywords: Artificial Intelligence, Starch Classification, Convolutional Neural Networks, Pharmaceutical Excipients, Microscopy, Transfer Learning

1.0 Introduction

Starch is a commonly used excipients in Pharmaceutical drug formulations, valued for its binding, disintegrating, and bulking properties [1, 2, 3]. Starch is one of the most widely and easily available polymers present in nature and can be obtained from various fruits and vegetables. It is also commonly used as food in our daily life [5]. However, starches from different botanical sources exhibit significant variability in granule morphology and physico-chemical behavior. Manual identification using microscopy is time-consuming and prone to error, especially when distinguishing morphologically similar starches [4]. Due to the different uses of starch obtained from different botanical origins, it is therefore necessary and important that identification and quality analysis is carried out on the starch sample during quality assurance and control, before and after incorporation into the drug formulation. These physico-chemical tests include organoleptic tests, bulk and tapped density, compressibility-index and Hausner ratio, angle of repose (USP 40- NF 35), solubility and swelling powers, amylose-amylopectin ratio, microscopy

such as Scanning Electron Microscopy (SEM), Fourier Transform and Infrared Spectroscopy (FT-IR) [6, 13] and chemical tests such as Toludiene test [7]. Methods and procedures often used in starch characterization are labour-intensive, time-consuming and require technical knowledge and skills available to just a few specialized individuals. Also, mis-identification of starches could occur due to human error because of their close physical similarities.

Machine Learning was defined by Samuel Arthur in 1959 as a field of study that gives computers the ability to learn without being explicitly programmed [8, 12]. It is also defined as A computer program designed to learn from experience **E** with respect to some task **T** and some performance measure **P**, if its performance on **T**, as measured by **P**, improves with experience **E**. [9]. Recent advances in Artificial Intelligence (AI) offer promising solutions for image-based classification tasks. Convolutional Neural Networks (CNNs) have demonstrated high accuracy in biomedical imaging, food quality control, and material sciences. In pharmaceutical formulations areas, AI has been applied to predict drug-excipient interactions, optimize formulations, and automate quality assurance [1]. CNNs are referred to as *state-of-the-art* for medical image analysis, demonstrating high accuracy in the detection of diseases like breast cancer and brain tumors from Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans [9,10,11]. These models often utilize transfer learning with architectures like ResNet to overcome the challenge of small annotated medical datasets [9].

This study explores the use of CNNs to classify starch granules from nine plant sources using microscopical images. By leveraging transfer learning with pretrained models, we aim to develop a robust, scalable tool for starch identification in Pharmaceutical drug formulation settings.

2.0 Materials and Methods

2.1 Starch Sample Preparation

Starches were extracted from nine botanical sources: maize, millet, cassava, potato, rice, wheat, oat, green peas, and tiger nut. Microscopic slides were prepared using glycerol and imaged at 640× magnification using a smart phone camera mounted on a light microscope.

2.2 Dataset

The dataset consists of 889 pictures from 9 different starches namely the starch of the kernel of Maize (*Zea mays*, Linn), the starch of fresh tubers of Cassava (*Manihot utilissima* Pohl.), the starch of root tuber of Irish Potato (*Solanum tuberosum* Linn.), starch of the grain of Wheat (*Triticum aestivum*), starch of the grain of Rice (*Oryza sativa*), the starch of the grain of Oat *(Avena sativa*), the starch of the peas of Green pea (*Pisum sativum* L.), the starch of the tubers of Tiger nut (*Cyperus esculentus* L.) and the starch of the grain of Millet (*Pennisetum glaucum* L.).

Table 1.**Names and quantities of different starch samples microscopic picture used**

2.3 Dataset and Pre-processing

A total of 889 images were collected and labeled. Images were re-sized to 224×224 pixels and normalized. Data augmentation included horizontal flipping and 10° rotation.

All the pictures were taken in ".jpeg" or ".heic" picture formats. The pictures in ".heic" were all then converted to ".jpeg" format using heic-to-jpg python package before any further processing was done on any of the images. Using torchvision, the images were all resized to 224 by 224 pixels to ensure the uniformity of all the pictures. The image values were also normalized with a mean of 0.5 and a standard deviation of 0.5 to reduce the skewness of the data and increase the speed of processing

2.4 Model Design Architecture

For the models, 2 pre-trained ResNet models (Fig. 2) were trained and evaluated, one pre-trained on Micronet dataset that contains over 100,000 microscopical images and the other was pre-trained on ImageNet dataset, which contain 14 million pictures of Ron random day to day activities . The models output layers were replaced by 3 fully connected layers having 500, 100 and 9 hidden nodes, ReLU activation function and 0.5 Dropout layer. The models were built with PyTorch .

Following the ImageNet competition, they receive as input images and consist of convolving filters (convolutions) that can extract features of pictures by applying convolutions defined by mathematical functions to multiple sections of an image to capture the local invariance and locality in images. They then apply activation functions to capture non-linear relationships in the image, and pooling functions to reduce the picture dimension space while retaining the most important information.

Two pretrained ResNet-18 models were used Fig. 2:

- **ImageNet-based model**: pretrained on general images
- **MicroNet-based model**: pretrained on microscopy-specific datasets

2.5 Training and Evaluation

Data was split into training (50%), testing (30%), and validation (20%) sets. Models were trained using Adam optimizer, early stopping, and evaluated using accuracy, precision, recall, and F1-score.

Upon successful training, the best performing models (based on the loss on the validation set) were then evaluated on the the held-out validation set using the precision, recall, accuracy and f1 score.

Precision given by equation 1 represents the number of images that the models correctly predicted their class compared to the number of all images the model predicted to be a particular class.

…..equation 1

Recall (also known as sensitivity) tells us the fraction of the correct labels/classes is captured by the predicted classes for the images, as shown in equation 2

…….equation 2

Accuracy tells the number of correctly predicted labels as a fraction of the total images available, shown in equation 3

….equation 3

F1 score is the fraction of the predicted value/class is actually correct (precision) and fraction of the actual value/class is correctly predicted and is shown in equation 4.

……equation 4

3.0 Results

3.1 Overall Performance

After evaluation of the models on the held-out validation set, the Micronet, ResNet model had an overall accuracy of 61.0% (Table 2) which is substantially lower than that obtained from the ImageNet ResNet model at 81.2%. The weighted precision and weighted recall of the Imagenet ResNet model 86% and 81.2% respectively were also higher than that obtained for the MicroNet ResNet model with 62% and 60% respectively shown in Table 2. The ImageNet-based model achieved an accuracy of 81.2%, representing a statistically significant improvement of 21.2% over the MicroNet-based model (60.0%) [95% CI: 13.7% to 28.7%; $p < 0.001$]. This robust performance gap confirms the superior adaptability of ImageNet features for light microscopy starch identification.

3.2 Class-Level Performance (ImageNet Model) Table 3.
3.3 Class-Level Performance (Micronet ResNet Model) Table 4.

4.0 Discussion

Data obtained from all analyses showed that the fine tuned pretrained ResNet model on the Imagenet dataset has a better performance than the fine tuned pretrained model on the Micronet dataset despite the greater similarity of the large-scale Micronet dataset to the dataset used for the research which both contain microscopic images as compared to the Imagenet dataset that contains random day to day things like dogs, cats, tables etc.

The concept of transfer learning (i.e. training a model initially on a large dataset and then fine tuning on a smaller dataset) was introduced to solve the insufficient dataset to train deep neural networks [14].

4.1 Confusion Matrix Insights

Millet and maize were frequently mis-classified due to morphological overlap. Cassava showed low precision but high recall, indicating over prediction. Table 5

It was difficult to compare other studies on starch classification with AI. On the individual performance on the different classes, pretrained ImageNet ResNet models significantly performed a lot better with precision values ranging from 0.70 (for Tigernut starch) to 1.00 (Green peas and Oat) except for a Cassava starch where the model has a precision of 0.41. The model also has significantly high recall values (> 0.85) for most of the starch samples except for Maize and Millet starch samples which were significantly low (0.23 and 0.31 respectively). However, it was observed that for the 3 classes in which the model didn't perform as great as others for either the precision or recall, a trade-off was made between the precision and recall i.e one of the two was low at a time, not both at the same time Table 2.

For the pretrained MicroNet ResNet models Table 3, very good precision values were obtained for Cassava starch, Irish potato starch, Rice starch and wheat starch (1.00, 1.00, 1.00 and 0.79 respectively) others gave a value ranging from 0.56 (Maize starch) to 0.12, Millet starch. The model also has high recall values (>0.7) for 5 out of the 9 different starch samples namely Irish potato starch (0.95), Maize starch (0.84), Rice starch (0.73), Tigernut starch (0.91) and Wheat starch (0.79). However, this model failed at identifying any of the millet starch samples having a precision value of 0.12 and recall value of 0.06. The ImageNet-based model outperformed MicroNet despite the latter being trained on microscopy images. This may be due to differences in imaging modality. MicroNet was trained on electron microscopy, while this study used light microscopy. Transfer learning proved effective, but performance was limited by dataset size and image quality.

Abbreviation	Meanings
AI	Artificial Intelligence
CNN	Convolutional Neural Network
FT-IR	Fourier Transform and Infrared Spectroscopy
MRI	Magnetic Resonance Imaging
CT	Computed Tomography
SEM	Scanning Electron Microscopy
USP-NF	United States Pharmacopoeia and The National Formulary
Abbreviation	Meaning
CONV	Convolutional Layer
Batch Norm	Batch Normalization
ReLu	Rectified Linear Unit
Max Pool	Maximum Pooling
Conv Block	Convolutional Block
ID Block	Identity Block
Avg Pool	Average Pooling
FC	Fully Connected Layer

Convolution Neural Networks (CNNs), the type of deep learning used in this research that has been previously pretrained on the Imagenet dataset have consistently shown the ability to demonstrate superior classification and segmentation tasks on microscopy [15,16,17], however given the dissimilarity between most images in the imagenet and most microscopic images, the MicroNet dataset was introduced, containing 100,000 microscopic images for pre-training models specifically for segmentation of the microscopical images as observed by [18].

5.0 Conclusion

The use of pretrained model on larger datasets can be used to augment small data sizes to get very good results. In addition for this task, while the dataset are microscopical images of starch samples obtained from the lab, model pretrained on Imagenet seem to perform a lot better than that pretrained on microscopical images from MicroNet. Future directions for this work will be to benchmark the accuracy, precision and recall of these AI models against human experts, which is the current global used method for identification of starch samples from their microscopic images given the domain knowledge required to be able to perform this task.

Another future direction for this project is the segmentation, identification and quantification of each individual starch sample in the microscopical image for purity estimation of the starch samples. This is because beyond the homogenous exhibition of the microscopical images, there exist the possibility of mix up of multiple and different starch samples inside the same microscopical image, thus the ability to segment and identify the different starch samples in the microscopical image and also estimate the quantity of this different starch samples will be of great use for the practical implementation of this idea.

Finally, AI-assisted starch classification using CNNs offers a scalable, objective tool for pharmaceutical quality control. The ImageNet model demonstrated superior performance, and future work will focus on enhancing accuracy and integrating this approach into industrial work flows.

Authors' contributions:

J.I.O and M.A: Conceptualization, Writing original draft. B.O, D.O.F and J.I: Formal analysis (equal), Methodology. J.I.O: Supervision,

All authors have read and agreed to the corrected and submitted version of the manuscript.

Acknowledgement:

We wish to thank Mr. Adeleke Tijani, Chief Technologist at the Department of Pharmacognosy Laboratory, University of Lagos. Nigeria for teaching us how to prepare the different starches.

Funding statement:

This research was self funded

Competing interests:

Authors declare no competing interest

References

[1] Ali M, et al. AI-driven formulation design in pharmaceutical excipients. J Pharm Anal. 2024.
[2] Chaksmithanont P, Bangsitthideth K, Arunprasert K, Patrojanasophon P, Pornpitchanarong C. Statistical-based optimization of modified *Mangifera indica* fruit starch as substituent for pharmaceutical tableting excipients. Polymers. 2024;16(18):2653. https://doi.org/10.3390/polym16182653
[3] Das S, Das MK, Jamatia T, Bhattacharya B, Mazumder R, Yadav PK, Mandal S. Advances of cassava starch-based composites in novel and conventional drug delivery systems: a state-of-the-art review. RSC Pharmaceutics. 2024;1(4):182–203. https://doi.org/10.1039/D3PM00008G
[4] Prajapati K, Patel R, Patel M. Maize starch. World J Pharm Res. 2025;14(9):261–74.
[5] Bordoloi BJ, Kalita B, Shil D. Properties, pharmaceutical application and various technique of chemical modification in native starch: a descriptive review. Int J Curr Pharm Res. 2019;11(4):54–9. https://doi.org/10.22159/ijcpr.2019v11i4.34950
[6] Azeh Y, Paiko Y, Abubakar A, Garba B, Ameh O. Isolation, modification and characterization of tiger-nut, maize, cassava and potato starch. J Appl Sci Environ Manage. 2018;22(5):713. https://doi.org/10.4314/jasem.v22i5.21
[7] Bisulca C, Odegaard N, Zimmt W. Testing for gums, starches, and mucilages in artifacts with o-toluidine. J Am Inst Conserv. 2016;55(4):217–27. https://doi.org/10.1080/01971360.2016.1239492
[8] Park T, Seong. Machine learning in biomedical engineering.Biomed Eng Lett. 2018;8(1):1–3. https://doi.org/10.1007/s13534-018-0058-3
[9] Mitchell TM. Machine learning.Vol. 1. New York: McGraw-Hill; 2007.
[10] Salehi AW, Khan S, Gupta G, Alabduallah BI, Almjally A, Alsolai H, Siddiqui T, Mellit A. A study of CNN and transfer learning in medical imaging: advantages, challenges, future scope. Sustainability. 2023;15(7):5930. https://doi.org/10.3390/su15075930
[11] Li M, Jiang Y, Zhang Y, Zhu H. Medical image analysis using deep learning algorithms. Front Public Health. 2023;11:1273253. https://doi.org/10.3389/fpubh.2023.1273253
[12] Kumar R, et al. Machine learning for excipient behavior prediction. Comput Biol Med. 2024.
[13] Shoukat R, Cappai M, Pilia L, Pia G. Rice starch chemistry, functional properties, and industrial applications: a review. Polymers. 2025;17(1):110. https://doi.org/10.3390/polym17010110
[14] Tan C, Sun F, Kong T, Zhang W, Yang C, Liu C. A survey on deep transfer learning. In: Artificial neural networks and machine learning – ICANN 2018. Springer Int Publ; 2018. p. 270–9. http://dx.doi.org/10.1007/978-3-030-01424-7_27
[15] Akers S, Kautz E, Trevino-Gavito A, Olszta M, Matthews B, Wang L, Du Y, Spurgeon S. Rapid and flexible semantic segmentation of electron microscopy data using few-shot machine learning. npj Comput Mater. 2022;8:107. https://doi.org/10.1038/s41524-022-00781-z
[16] Durmaz AR, Müller M, Lei B, Thomas A, Britz D, Holm EA, Eberl C, Mücklich F, Gumbsch P. A deep learning approach for complex microstructure inference. Nat Commun. 2021;12(1):6272. https://doi.org/10.1038/s41467-021-26565-5
[17] Goetz A, Durmaz AR, Müller M, Thomas A, Britz D, Kerfriden P, Eberl C. Addressing materials' microstructure diversity using transfer learning. NPJ Comput Mater. 2022;8(1). https://doi.org/10.1038/s41524-022-00703-z [18] Stuckner J, et al. Micro structure segmentation with deep learning encoders.NPJ Comput Mater. 2022.

GRAPHICAL ABSTRACT

DEEP LEARNING-BASED CLASSIFICATION OF PHARMACEUTICAL STARCHES USING MICROSCOPY AND TRANSFER LEARNING

The research highlights for the research article titled: Deep Learning-Based Classification of Pharmaceutical Starches Using Microscopy and Transfer Learning

High accuracy: the Artificial Intelligence [AI] model helped to achieve high accuracy in classifying microscopic starch images, with accuracies ranging from 61.0% to 81.2%. This is far better than the accuracy of human experts.

Speed: the AI model classified microscopic starch images of 9 different samples much faster than any human experts though not directly studied here thus saving significant time and resources.

Robustness: This AI model is robust because the different variations in microscopic starch images were noted, such as different magnifications, lighting conditions, and sample preparation techniques. This makes it suitable for use in a variety of real-world settings and specifically for resource poor situations as found in sub-Saharan Africa.

Table 1. Names and quantities of different starch samples microscopic picture used

S/N	Starch Sample	Microscopical Pictures Quantity
1	Cassava	110
2	Green peas	119
3	Irish Potato	100
4	Maize	100
5	Millet	103
6	Oat	81
7	Rice	91
8	Wheat	112
9	Tigernut	83
	Total	**889**

Table 2: Data on accuracy, precision and recall of the models used

Model	Accuracy (%)	Precision	Recall	F1-score
ImageNet-CNN	81.2	0.86	0.81	0.77
MicroNet-CNN	60.0	0.62	0.60	0.58

Table 3: Result of Imagenet ResNet Model showing the Precision, Recall and f1-score

Starch sample	Precision	Recall	f1_score	Support	Labels
Cassava	0.411765	0.933333	0.571429	15.0	Imagenet
Green peas	1.000000	1.000000	1.000000	26.0	Imagenet
Irish potato	0.916667	1.000000	0.956522	22.0	Imagenet
Maize	0.857143	0.230769	0.363636	26.0	Imagenet
Millet	0.833333	0.312500	0.454545	16.0	Imagenet
Oat	1.000000	0.937500	0.967742	16.0	Imagenet

Table 4: Result of pretrained Micronet ResNet Model showing the Precision, Recall and f1-scores

Starch sample	Precision	recall	f1_score	support	labels
Cassava	1.000000	0.266667	0.421053	15.0	micronet
Green peas	0.500000	0.461538	0.480000	26.0	micronet
Irish potato	1.000000	0.954545	0.976744	22.0	micronet
Maize	0.564103	0.846154	0.676923	26.0	micronet
Millet	0.125000	0.062500	0.083333	16.0	micronet
Oat	0.285714	0.125000	0.173913	16.0	micronet
Rice	1.000000	0.733333	0.846154	15.0	micronet
Tigernut	0.454545	0.909091	0.606061	22.0	micronet
Wheat	0.789474	0.789474	0.789474	19.0	micronet

Table 5: Confusion matrix of pretrained MicroNet model with actual class as the column and predicted classes as the row

Predicted/ Actual	Cassava	Green peas	Irish potato	Maize	Millet	Oat	Rice	Tigernut	Wheat
Cassava	4	6	0	3	0	0	0	2	0
Green peas	0	12	0	0	6	0	0	8	0
Irish potato	0	0	21	1	0	0	0	0	0
Maize	0	3	0	22	0	1	0	0	0
Millet	0	1	0	0	1	0	0	14	0
Oat	0	1	0	13	0	2	0	0	0
Rice	0	0	0	0	0	0	11	0	4
Tigernut	0	1	0	0	1	0	0	20	0
Wheat	0	0	0	0	0	4	0	0	15

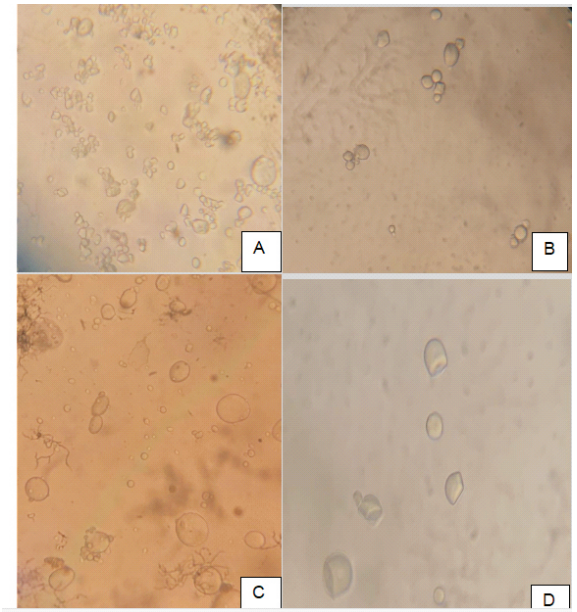
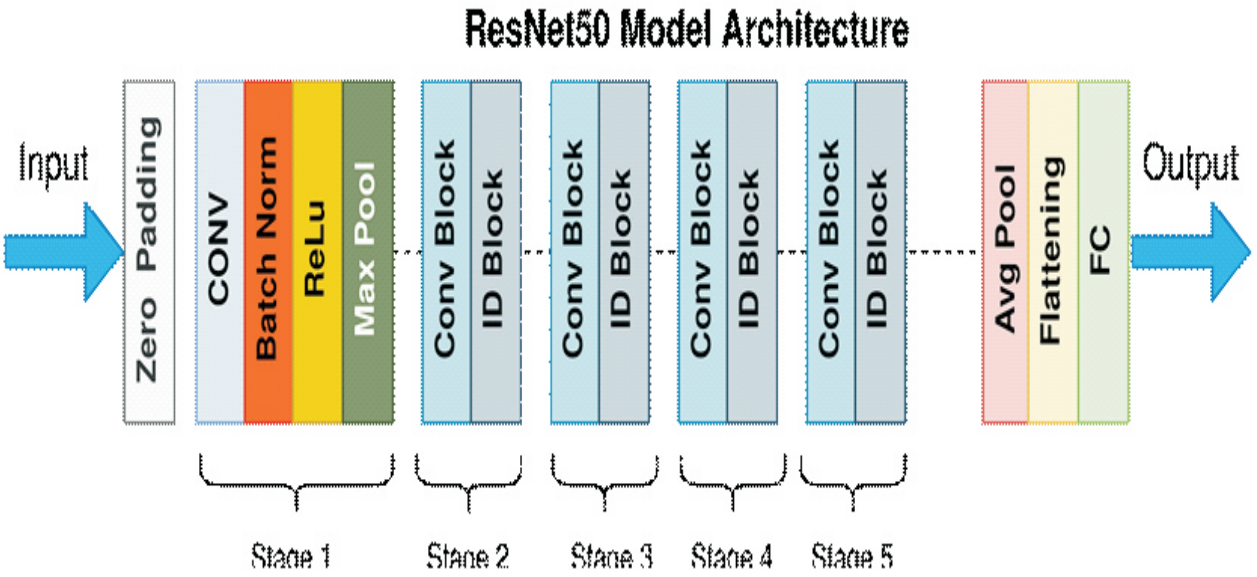


Fig 1. Examples of microscopical images of the starches: from left to right,[A]: Maize starch image [B]: Oat starch image [C] Wheat starch image [D] Cassava starch image *Mag X640*



CONV: The core building block of a CNN that applies filters to the input to create feature maps. **Batch Norm**: A technique to standardize the inputs to a layer, which helps stabilize and speed up the training of deep neural networks.**ReLu**: An activation function that introduces non-linearity into the model by outputting the input directly if it is positive, and zero otherwise. **Max Pool**: A sample-based discretionary process that down-samples the input representation to reduce its dimensional and computational cost. **Omv. Lock:** A sequence of layers (typically CONV, Batch Norm, and ReLu) grouped together as a functional unit in the ResNet architecture. **ID Block**: A standard residual block used in ResNet when the input and output dimensions are the same, allowing the model to learn identity mappings. **Avg Pool**: A pooling operation that calculates the average value for patches of a feature map, used here before the final output stages. FC: Fully connected Layer that connects every neuron in one layer to every neuron in another, typically used at the end of the network for final classification.

Fig 2. ResNet18 model architecture