Open-source microscopic solution for classification of biological samples

Robert Archibald[a], Graham M. Gibson[a], Samuel Westlake[b], and Akhil Kallepalli[a,*]

[a]School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, United Kingdom
[b]Centre for Electronic Warfare, Information and Cyber, Cranfield University, Shrivenham, Swindon SN6 8LA, United Kingdom

ABSTRACT

Open-source technologies and solutions have paved the way for making science accessible the world over. Motivated to contribute to the direction of open-source methods, our current research presents a complete workflow of building a microscope using 3D printing and easily accessible optical components to collect images of biological samples. Further, these images are classified using machine learning algorithms to illustrate both the effectiveness of this method and show the disadvantages of classifying images that are visually similar. The second outcome of this research is an openly accessible dataset of the images collected, OPEN-BIOset, and made available to the machine learning community for future research.

The research adopts the OpenFlexure Delta Stage microscope (https://openflexure.org/) that allows motorised control and maximum stability of the samples when imaging. A Raspberry Pi camera is used for imaging the samples in a transmission-based illumination setup. The imaging data collected is catalogued and organised for classification using TensorFlow. Using visual interpretation, we have created subsets from amongst the samples to experiment for the best classification results. We found that by removing similar samples, the categorical accuracy achieved was 99.9% and 99.59% for the training and testing sets. Our research shows evidence of the efficacy of open source tools and methods. Future approaches will use improved resolution images for classification and other modalities of microscopy will be realised based on the OpenFlexure microscope.

Keywords: OPEN-BIOset, open source, additive manufacturing, microscopy, machine learning

1 INTRODUCTION

Microscopy is an essential component of histology and a valuable tool for healthcare and medical research. In disease diagnosis, microscopic examination of thin-film blood smears is the gold standard for diagnosing many blood-borne diseases such as malaria,[1] and biopsies remain the principal method of cancer diagnosis.[2] Meanwhile, in the field of medical research, microscopic examination of common model organisms–such as fruit flies and zebra fish–is integral to the study of human brain degenerative diseases, human embryonic development, and host-pathogen interaction.[3,4] Microscopy is also fundamental to various other scientific fields, such as forensic science, where it is key to the identification of trace evidence, such as fibres, fragments, and fingerprints.[5–7]

Though an important tool, laboratory grade microscopes are prohibitively expensive to many and their maintenance and operation often demands skilled technicians and microscopists. Consequently, access to microscope capabilities is often unavailable for purposes such as biological research, health care, and education. For example, in rural areas, access to quality instrumentation and qualified microscopists can significantly limit the accurate diagnosis of diseases, such as malaria.[8] In education, a lack of access to microscopes and STEM outreach programmes strengthens the stereotypes that STEM is boring, irrelevant and difficult, thus reducing the participation of the next generation in furthering humanity.[9]

Furthermore, in Ethiopia, malaria is a huge problem and requires the gold standard of microscopy to analyse. Mitiku *et al.* finds that due to a lack of expert microscopic analysis identification of malaria in a blood film examination was low and that consistent supervision and training are indispensable in the fight against malaria.[10] Consequently, improved access to microscope instrumentation and professional microscopic analysis will significantly enhance healthcare standards by aiding medical research and accelerating the diagnosis of disease.

* For additional information regarding the research and the open-source dataset, please send all correspondence to A.K.
E-mail: Akhil.Kallepalli@glasgow.ac.uk

Frontiers in Biophotonics and Imaging, edited by Sumeet Mahajan, Stefanie Reichelt, Proc. of SPIE
Vol. 11879, 118790D · © 2021 SPIE · CCC code: 0277-786X/21/$21 · doi: 10.1117/12.2599435

To deliver such improvements, high-quality microscopes must become more accessible and simpler to maintain. With the development of 3D printing technology, the OpenFlexure team have overcome some of these problems. There are certain aspects of a microscope that will always require daily maintenance such as the lenses and objectives, where keeping dirt and dust out without scratching or damaging the apparatus is paramount. However, general 'wear and tear' of microscopes can be overcome with the OpenFlexure microscopes since part replacement simply requires a 3D printer, even significant damage to the microscope would not be too much of a financial burden to a research team with the complete assembly of an OpenFlexure delta stage model costing less than £200.[11]

To fully exploit the benefits of accessible microscopes, a dramatic increase in the availability of expert microscopic analysis is required. This can be achieved through the development of autonomous, intelligent microscopes which are capable of minimising the workload of microscopists, thus facilitating an increase in productivity. The potential of such systems are being demonstrated across a wide range of domains. For example, in the diagnosis of malaria manual light microscopy of thin-film blood smears is the gold standard,[1] yet this process relies on the identification of *plasmodim falciparum* which often may only be visible in fewer than 1 in 1,000 erythrocytes (red blood cells). As a result, accurate malaria diagnosis can be a time-consuming process, however newly emerging computer vision algorithms have demonstrated the ability to detect infected erythrocytes with human-level sensitivity,[8] thus considerably decreasing the time to diagnosis. Similar systems under development include the rapid diagnosis of COVID in chest X-ray images,[12] heart disease from clinical observations,[13] and cancer from CT scans.[14]

In addition to healthcare and medical research, intelligent autonomous microscopes are also demonstrating potential as powerful educational tools. In this paper, we initiate the development of an intelligent computer vision algorithm for use in conjunction with an easily accessible microscope – the 3D printable OpenFlexure Delta Stage.[15] This is a low-maintenance, open-source and versatile microscope, capable of fluorescence and transmission microscopy. This microscope can also be driven autonomously and is fully motorised. A computer vision algorithm was developed to classify the contents of various microscope slides depicting an assortment of animal and plant tissue.

Autonomous microscope image analysis, however, constitutes a challenging computer vision task which requires a highly generalised model of the complex relationships between pixels. Current methods of generating such models typically rely on Convolutional Neural Networks (CNNs), a type of artificial neural network which are capable of efficiently modelling abstract visual concepts. However, the optimisation and validation of CNNs requires large sets of annotated example data which is comprehensively representative of the test conditions. Consequently, this paper also presents the OPEN-BIOset,[16] a dataset of microscope imagery that we make openly accessible for any machine learning applications and algorithms. This dataset, designed for image classification tasks, was collected using a 3D printed delta stage OpenFlexure microscope and depicts 19 different animal and plant tissue samples.

The contributions of this paper can be summarised as: 1) OPEN-BIOset,[16] an annotated image dataset of animal and plant tissue for the development of autonomous intelligent microscopes and, 2) an analysis of the classification performance of a modern ConvNet with respect to OPEN-BIOset.

2 BACKGROUND

The main component of this research is the 3D printed microscope. The OpenFlexure microscope is made available online in three versions. We chose the advanced version known as the delta stage microscope that provides the advantage of having a fixed optics module.[17,18] The OpenFlexure team, in Collins *et al.*,[17] discuss the benefits of a one-piece flexure stage and the microscopes imaging capabilities. They show that the microscope has versatile imaging capabilities such as trans-and epi-illumination, polarisation contrast imaging, and epi-florescence imaging. The OpenFlexure microscope has also been adapted to perform super resolution imaging. By applying the super resolution radial fluctuations algorithm, Patton *et al.*[19] obtained images with a resolution of 115 nm, which is beyond the diffraction limit. The delta stage microscope was adopted and extended by Vallejo Ramirez *et al.*[20] to develop OptiJ. An Optical Projection Tomography (OPT) system is capable of imaging large transparent specimens up to 13 mm tall and 8 mm deep. OPT is used to image samples in 3D and Vallejo Ramirez *et al.*[20] share the open-source hardware and software design with the intention to further develop this software, utilising the open-source platform.

The first artificial neural network, named Perceptron, was created for pattern recognition.[21] A noticeable upgrade in artificial neural networks was made with Neocognitron.[22] The first modern Convolutional Neural Network (CNN) was designed by Yann LeCun – the LeNet,[23] which used the MNIST dataset for training. Subsequently, projects such as

Google Brain[24] (which used a deep neural network to detect patterns in images and videos) and AlexNet[25] (which created ReLU[26] as an activation function for CNNs[27–30] and used for image analysis to vastly improve their efficiency) have been supported, researched and found application in real world problems. Artificial intelligence and machine learning have also found applications in cyber security,[31] autonomous driving[32] and cancer research.[33]

Google Brain built Tensorflow as an API for implementing and executing machine learning algorithms. TensorFlow supports both C++ and Python as front-end languages. This interface is open-source, allowing many users to utilise, adopt and share their work. This has created huge advancements in machine learning applications and development.[24,34] At the University of Toronto, Krizhevsky *et al.*[25] developed a deep CNN consisting of five convolutional layers, multiple max-pooling layers and three fully-connected layers with a softmax output layer totalling 60 million parameters. They classified 1.2 million images in the ImageNet LSVRC-2010 contest into 1000 different classes. They achieved record breaking results using such a large dataset with a top-5 error rate of 15.3% for their testing dataset, significantly better than the nearest competitors. A dataset of potato tuber was established using a microscope. Bismas *et al.*[35] created a dataset consisting of 9811 unstained and 6127 stained potato samples. Through deep learning assessment, by utilising a CNN, and using a train-test split of 80-20, an accuracy of 94.27% was achieved for the unstained dataset.

Machine learning systems can process massive amounts of data very quickly and learn in real-time, improving performance on tasks such as classification. By creating systems that mimic neurons in the brain, multi-layer perceptrons[21] can process an input layer, such as an image, and produce an output layer, likely a probability. In-between the input layer and output layer are hidden layers. These hidden layers, specifically convolutional layers used for this CNN, back-propagate, training the algorithm to reduce errors and improve classification accuracy.[27–30]

CNNs can process images and classify them into categories. For example, CIFAR-10[36] is a dataset consisting of 60,000 images and 10 classes, each with 6000 images. Applying this dataset to a CNN you can test your algorithms capability at categorising images correctly and the rate at which it learns. Section 5.2 discusses the CNN applied to this biological dataset in more detail.

3 IMPLEMENTATION

The primary aim of this project is to illustrate the advantages of a completely open-source workflow using an affordable microscope. We set about building the motorised OpenFlexure delta stage microscope to acquire images from biological samples. The research subsequently produced an open-access dataset named OPEN-BIOset. For the dataset, we were aware of two challenges:

- The dataset must be large enough to prevent over-fitting of the model. Essentially, if the training subset of data becomes too familiar with the features of the images, then the algorithms performance is negatively impacted.
- The dataset must have enough distinguishable features for a classification algorithm to work.

3.1 Microscope

The setup and construction of the microscope is adopted from OpenFlexure (https://openflexure.org). The delta stage was chosen for its better control and fully motorised design. The microscope optics includes a x40 RMS threaded microscope objective lens for imaging the sample. A white LED, a condenser lens (f = 7.5 mm, diameter = 13 mm), and a diffuser plate are used for the illumination module. A fixed optics module achieves a 12 mm x 12 mm x 4 mm field of view using stepper motors controlled with a motor controller board connected to a Raspberry Pi, to move the sample slide holder. The images were captured using a 5 MP Raspberry Pi camera with images being saved at 2560 x 1440 pixel resolution. All printed parts were black PLA (Polylactic Acid) to reduce stray light around the microscope as seen in Figure 1. The delta stage microscope consists of three main parts, described in the next subsection.

3.1.1 Actuators & motors

A 3D printed one-piece body with three actuators is the work horse of the microscope. Equipped with a 12x12x4 mm field of view and a fixed optics module, the delta stage is an upgraded edition of the original OpenFlexure microscope.[17,18] Connecting three stepper motors (32 steps per revolution with a x64 gearing ratio, 2048 total steps per revolution) via a motor controller board, allows for X, Y, Z–directional control of the sample. The motor controller board consists of an Arduino Nano connected to a Raspberry Pi.

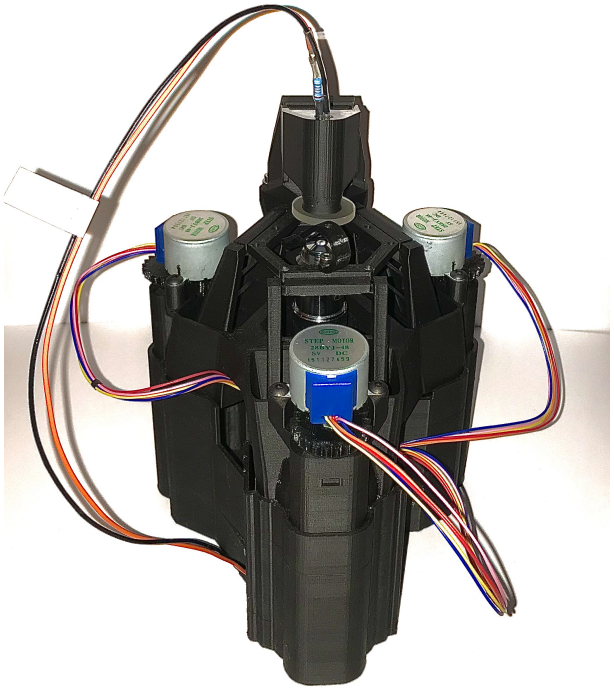

Figure 1: The fully motorised 3D OpenFlexure delta stage microscope.

3.1.2 Optics

The fixed optics module consists of a x40 RMS threaded microscope objective lens, a tube length correction lens (f = 50 mm, diameter = 12.7 mm) and a Raspberry Pi camera V1.3, capable of 5MP resolution (the original Pi camera lens has been removed). There is a white LED, a condenser lens (f = 7.5 mm, diameter = 13 mm) and a diffuser plate for the illumination module. This same configuration was used for all the data collected in this work.

3.1.3 Raspberry Pi & software

A Raspberry Pi model 2B with the software from OpenFlexure was used to control the motors and take images (2560 x 1440 pixel resolution).

3.2 Open-source Dataset

The goal of this project was to implement a machine learning-based classification approach to microscope images taken from an open-source microscope. The dataset is required to be both large enough to avoid over-fitting while containing distinct enough image characteristics such that the classification algorithm achieves a clear output. The delta stage microscope produced over 1900 images (100 different images from each of the 19 microscope sample slides). Figure 2 shows a random collection of each sample's images. Visually, it is noted that many samples are distinct but some share similar characteristics.

The complete dataset (listed in Table 1) consisted of backgrounds, coprimus mushroom set, dense connective tissue, dog cardiac muscle, dog oesophagus, dog skeletal muscle, house bee mouth parts, hydrilla verticillata leaf, lillium ovary, nymphaea of aqustio stem, onion epidermis, pig motor nerve, pine leaf, pine stem, pumpkin stem, rabbit spinal cord, sunflower stem, tilia stem, young root of broad bean and zea stem samples. From this, five datasets were produced. Dataset 1 consisted of all samples, Dataset 2 consisted of all animal samples, Dataset 3 consisted of all plant samples and Datasets 4 and 5 were refined datasets established after analysis of the confusion matrices.

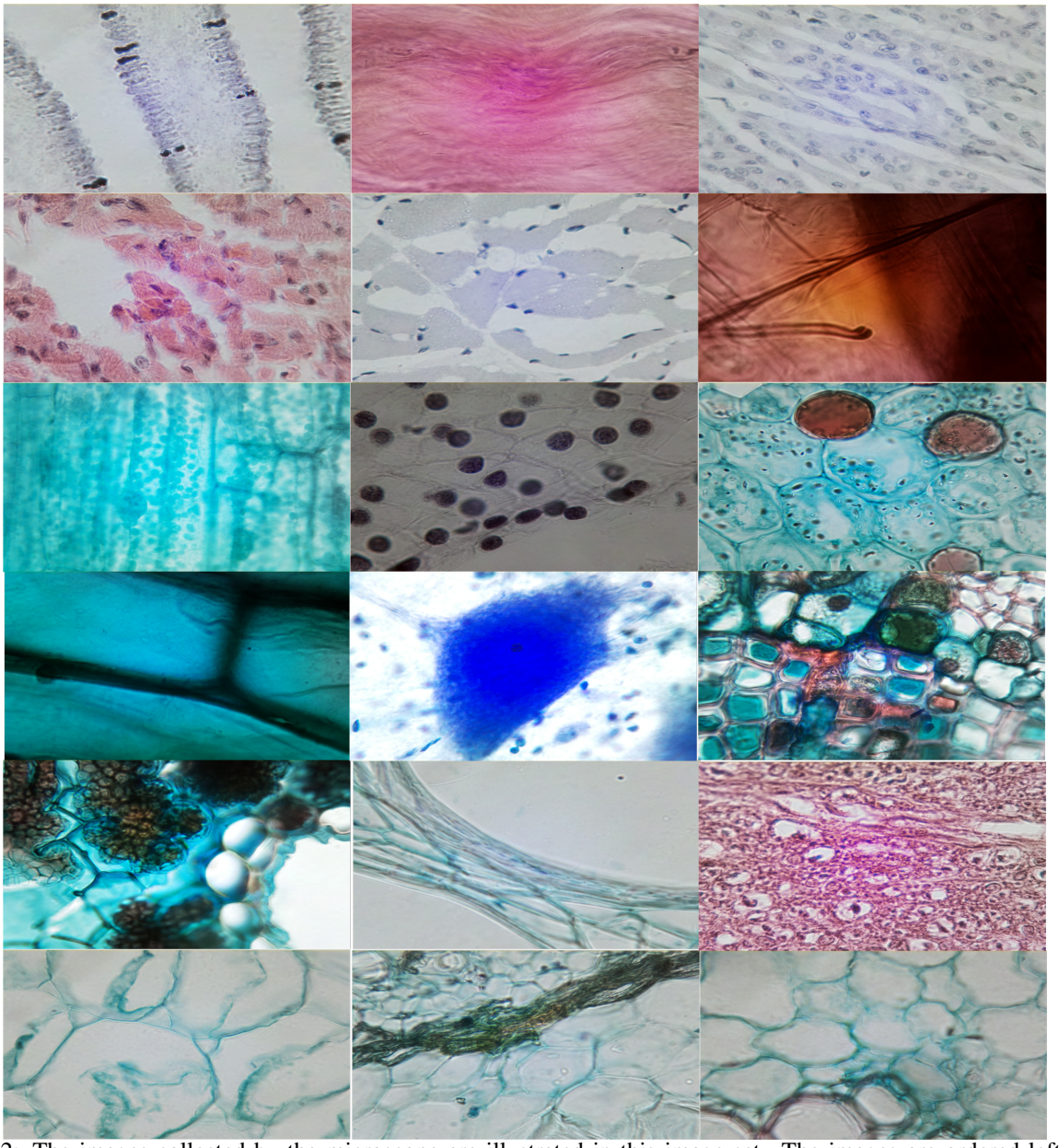

Figure 2: The images collected by the microscope are illustrated in this image set. The images are ordered left-right, top-bottom in the same sequence as shown in Table 1 except "Zea Stem" (omitted for image set alignment).

3.3 Classification Algorithm

These datasets were then applied to the multi-layered convolutional neural network (CNN). Figure 3 illustrates the classification algorithm's architecture. The CNN consisted of two convolutional layers plus ReLU activation with two max-pooling layers succeeding the convolutional layers. There was then a flatten layer with two dense layers and a softmax activation as the output.

Five datasets were applied to the classification algorithm. Dataset 1 is the complete dataset consisting of all the slides. Dataset 2 contains only animal slides and dataset 3 contains only plant slides. Analysing the confusion matrices for datasets 1-3 decided what samples should be included in dataset 4 and dataset 5.

100 images in different areas of each sample slide were taken in order to cover as much of the sample as possible.

Table 1: The samples considered in each individual dataset are listed below. Beginning with Dataset 1 that contains all the samples, similar images are selectively removed through to Dataset 5 (which considers a subset of the samples). This screening process is necessary in this study.

Biological sample	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5
Backgrounds	✓	✓	✓	✓	✓
Coprimus Mushroom Set	✓	✗	✓	✓	✓
Dense Connective Tissue	✓	✓	✗	✓	✓
Dog Cardiac Muscle	✓	✓	✗	✓	✗
Dog Esophagus	✓	✓	✗	✓	✓
Dog Skeletal Muscle	✓	✓	✗	✓	✓
House Bee Mouth Parts	✓	✓	✗	✓	✓
Hydrilla Verticillata Leaf	✓	✗	✓	✓	✓
Lillium Ovary	✓	✗	✓	✓	✓
Nymphaea of Aqustio Stem	✓	✗	✓	✗	✗
Onion Epidermis	✓	✗	✓	✓	✓
Pig Motor Nerve	✓	✓	✗	✓	✓
Pine Leaf	✓	✗	✓	✗	✗
Pine Stem	✓	✗	✓	✓	✓
Pumpkin Stem	✓	✗	✓	✗	✗
Rabbit Spinal Cord	✓	✓	✗	✓	✓
Sunflower Stem	✓	✗	✓	✗	✗
Tilia Stem	✓	✗	✓	✗	✗
Young Root of Broad Bean	✓	✗	✓	✓	✓
Zea Stem	✓	✗	✓	✗	✗

Exposure settings values ranged from as low as 476 J/m^2 for the "Coprimus Mushroom Set" sample and as high as 24,969 J/m^2 for the "Onion Epidermis" sample and the exposure settings were adjusted depending on the thickness of the samples. The dataset was organised into a folder "imageslides" consisting of a sub-folder for each sample slide and a sub-folder for background images. This was then directly applied to the classification algorithm with sub-folders removed depending on what dataset was being applied. Table 1 shows which slides were used in each of the datasets.

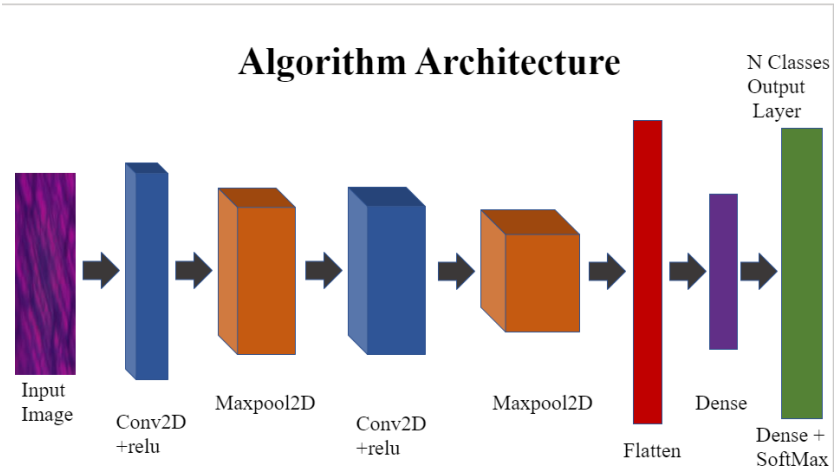

Figure 3: Architecture of the Convolutional Neural Network.

A multi-layered CNN was used as the classification algorithm. Utilising Tensorflow's API Keras, in Python as a sequen-

tial model allowed us to pass images through layer by layer. Figure 3 is a representation of the algorithm architecture. The two convolutional layers (Conv2D, Figure 3) deal with the input images, which appear as 2-dimensional matrices.[25,27–30] These layers identify characteristics in the samples and will recognise them in different positions, meaning the model can learn rapidly. Setting the number of nodes in the first layer to be 128 and 256 in the second layer, respectively, worked best for this model. A ReLU (Rectified Linear Activation Unit) activation function accompanies the Conv2D layers and improves the learning rate and output values.[26,37] The ReLU activation function is the superior activation unit and supersedes the sigmoid and tanh activation functions which were the default activation functions for a CNN.[37]

Max-pooling layers (Maxpool2D) are inserted between the Conv2D layers.[25,27,28] These layers simplify the information produced from the Conv2D layers and produces a condensed collection of the information. Stating (2, 2) for pool size and strides in the Maxpool2D layers keeps the maximum values in each 2 x 2 pool, creating a more efficient model. It is important to note that the max-pooling layers contain the same number of nodes (or filters) as the Conv2D layers they subsequently follow. In order to have a softmax output, the layer must have a "Flatten" layer to change the data from a 3D tensor to a 1D tensor.[25] The model then uses a "Dense" layer to change the dimensions of the vector and a "Dense + SoftMax" layer is the final layer. The softmax layer can classify more than two samples and produces the output values as probabilities by taking the 'evidence' learnt from the Conv2D layers, i.e whether an image belongs to a particular sample. The total number of trainable parameters for this model is 3,986,451, varying slightly depending on the dataset applied.

The optimiser applied to this CNN is the Adam algorithm.[38] Adam is a stochastic optimisation method (using random variables to optimise an algorithm) designed to work quickly and efficiently with large amounts of parameters. The default learning rate (0.001) was used in the CNN. Categorical cross-entropy was the loss function in the algorithm. This computes the categorical cross-entropy loss between the training and validation subsets for multiple classifications. This works by measuring the difference between two probability distributions.

The data is split into testing and training (validation) subsets with an 80-20 split and applied to an image data generator. The image data generator normalises all images for both test-train subsets as well as vertically flipping images, rotating images $\pm 15^{\circ}$ and shifting images $\pm 5\%$ up, down, left and right in real-time in the testing subset.

4 RESULTS AND DISCUSSIONS

The AUC (area under the curve) output values represent the probability of correctly categorising an image.[39] As a function of epochs, the calculated AUC values are plotted for each dataset analysed using the CNN. Secondly, the categorical accuracy is a metric output that gives the percentage of each epoch's categorisation correctness. The percentages are plotted against epochs for each datasets train-test subsets. A confusion matrix for each dataset is produced directly from the CNN. A summary of the results is shared in Table 2

Table 2: Summary of results from all datasets assessed in this study.

Peak Values	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5
AUC Training	0.9998	0.9967	0.9997	0.9996	1.0
AUC Validation	0.9965	0.9984	0.9974	0.999	1.0
Classification Accuracy Training	97.24%	92.68%	97.09%	96.45%	99.9%
Classification Accuracy Validation	93.96%	95.71%	92.95%	95.02%	99.59%

4.1 Dataset 1

Dataset 1 consisted of all nineteen sample slides. Applying dataset 1 to the model, a peak classification accuracy for the training and validation subsets achieved was 97.24% and 93.96%, respectively. This indicates a highly accurate CNN. Plotting the AUC values for the test-train data against the epochs (Figure 4) shows a very consistent curve with minimal fluctuation. A rapid increase from an initial value of 0.6681 for the training subset to 0.9916 in just 8 epochs indicates that the model learns quickly and efficiently, within the scope of this study and the samples imaged.

The categorical accuracy shows a rapid initial increase in values from 11.05% and 18.64% to 82.97% and 82.15% in just 13 epochs for the training and validation subsets, respectively. The CNN shows a fluctuating increase in values for both the training and testing subsets with a peak categorical accuracy of 97.24% and 93.96%, respectively.

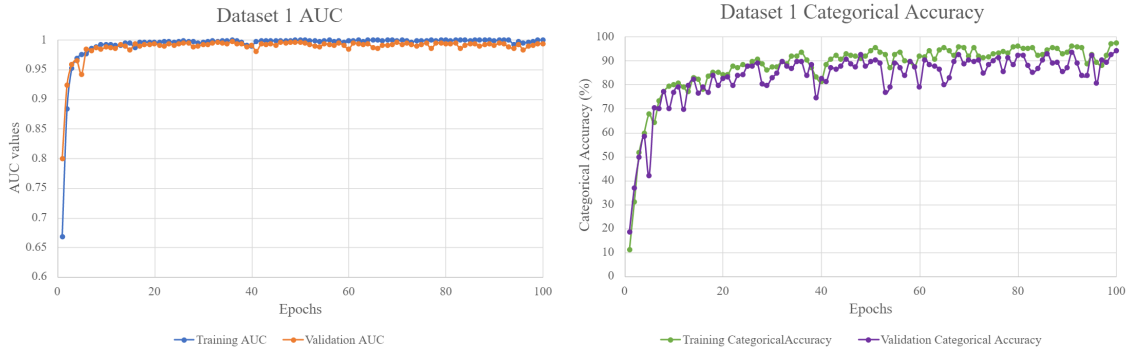

Figure 4: AUC and categorical accuracy illustrates a steep increase after less than 15 epochs for Dataset 1. This is indicative of a quick and efficient learning phase of the CNN.

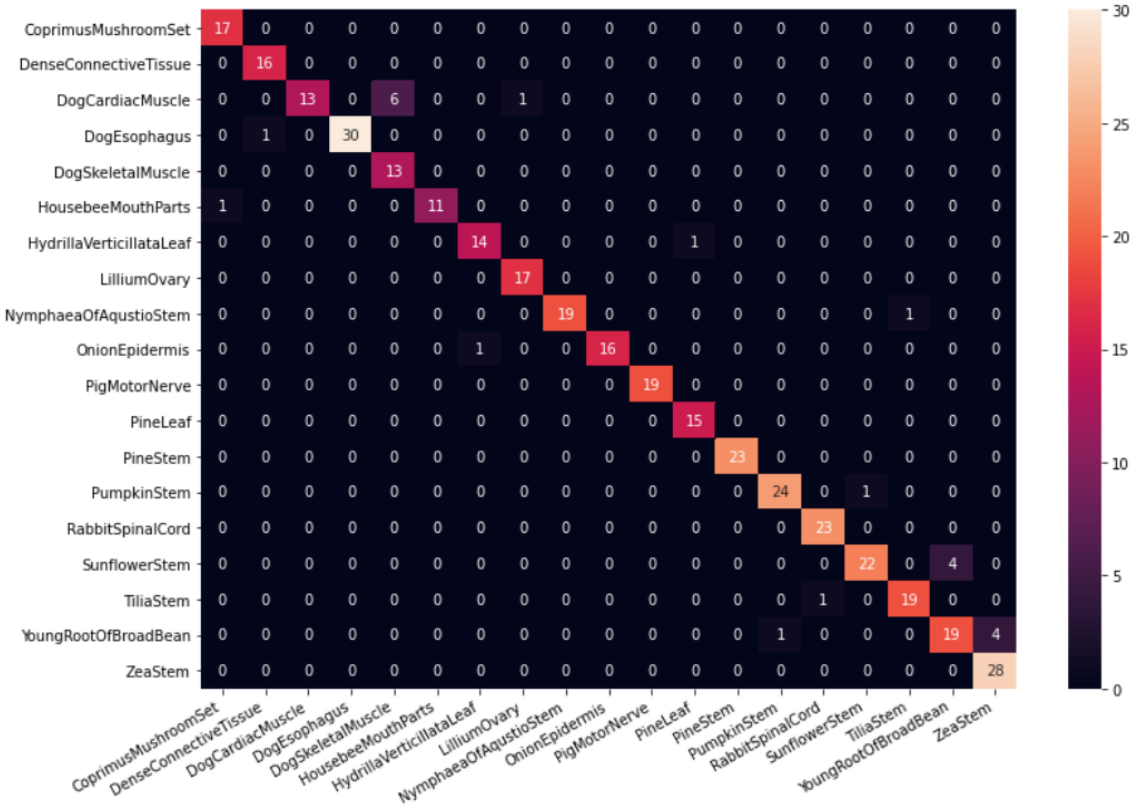

Figure 5: The confusion matrix for Dataset 1 assists in identifying images that are difficult to decipher during the validation stage.

In combination with these values (Figure 4), assessing the confusion matrix (Figure 5) reveals that the "Dog Cardiac Muscle" and the "Dog Skeletal Muscle" result in six false positive results. Another conflict was observed for "Zea Stem" and "Sunflower Stem", with the "Young Root of Broad Bean" sample with four false positive results each.

4.2 Dataset 2

The dataset 2 consists of all seven animal samples (refer to Table 1 for details). This produced a peak classification accuracy of 92.68% and 95.71% for the training and validation subsets, respectively. The reduced training classification accuracy is attributed to a smaller data set so over-fitting is highly probable. Plotting the AUC values for the test-train data against the

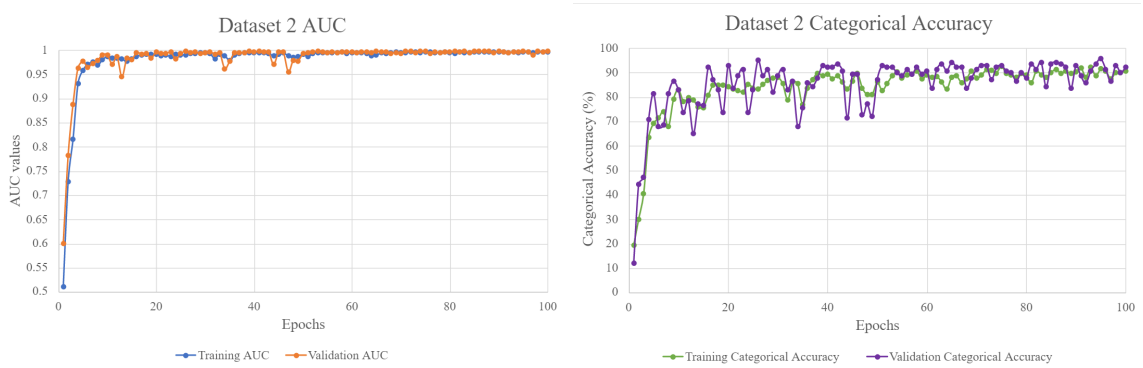

Figure 6: AUC and categorical accuracy plots illustrate the inferences made regarding over-fitting of data due to fewer samples being input to the CNN when using Dataset 2.

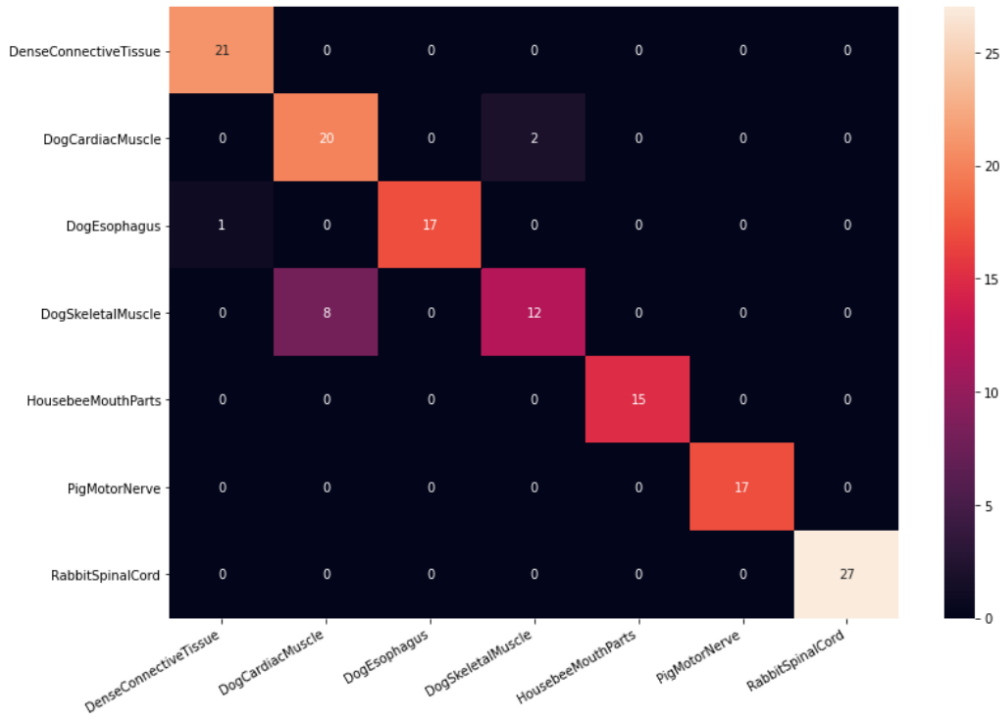

Figure 7: The confusion matrix illustrates the false positives when adding Dataset 2 to the CNN, i.e. mis-categorisations between “Dog Cardiac Muscle” and “Dog Skeletal Muscle.”

epochs (Figure 6) shows a sharp rise in probability of correct categorisation and the training data shows a consistently high AUC curve. The validation AUC shows some fluctuation before 50 epochs and then consistently outputs a high probability with a peak AUC of 0.9967 and 0.9984 for the training and validation subsets, respectively. Plotting categorical accuracy as a percentage against epochs shows the validation subset fluctuates with jumps over 10% before 50 epochs and then matches the training subset.

The confusion matrix (Figure 7) reveals that the “Dog Cardiac Muscle” and the “Dog Skeletal Muscle” introduce conflict during classification with two false positive results and eight false negative results. The other samples in dataset 2 show very little confusion between each other.

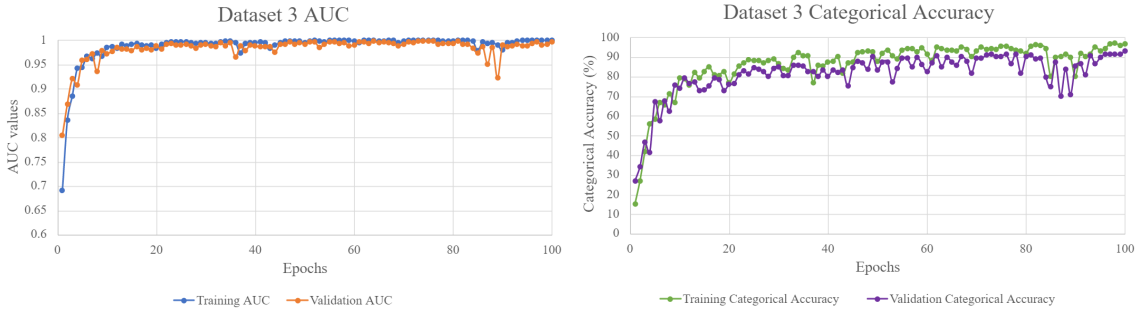

Figure 8: AUC and categorical accuracy graphs illustrate the results obtained when assessing the Dataset 3 with the CNN workflow.

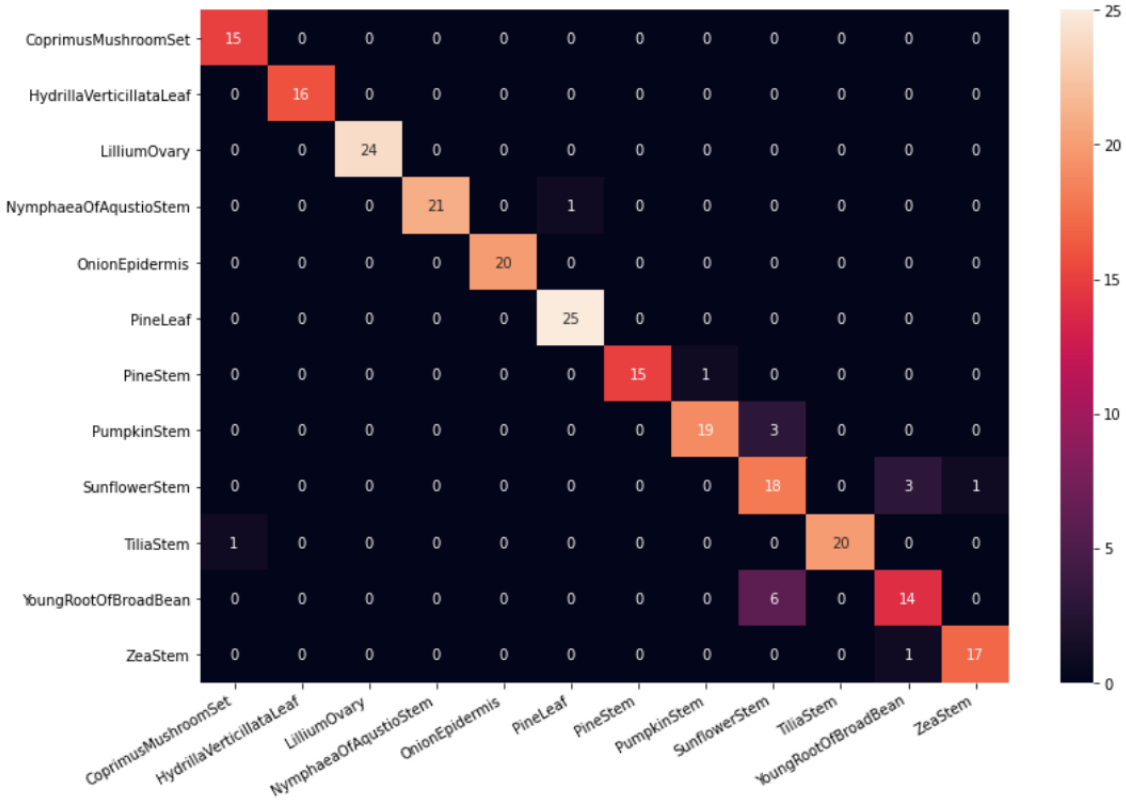

Figure 9: The confusion matrix for Dataset 3 illustrates the false positives when comparing the “Sunflower Stem” with “Pumpkin Stem” and “Young Root of Broad Bean”. The other sample images were accurately identified.

4.3 Dataset 3

In order to compare the plant samples in the combined dataset, dataset 3 consisted of all twelve plant samples (refer to Table 1 for details). Running this dataset through the model produced a peak classification accuracy of 97.09% and 92.95% for the training and validation subsets, respectively. Plotting the AUC values for the test-train data against epochs (Figure 8) showed the dataset mirrors the behaviour with the validation AUC values being consistently lower. There is also a decrease for the validation AUC values at epoch 87 with a value of 0.9508 and epoch 89 with a value of 0.9216. We conclude that this is due to a random batch of images containing multiple images from samples that cannot be deciphered by the algorithm.

Plotting categorical accuracy shows that the test-train subsets fluctuate but follow a similar upwards trend. We obtain

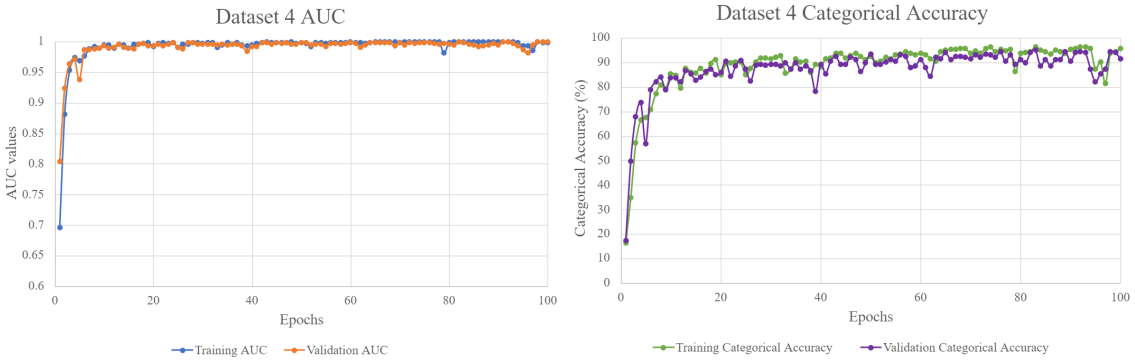

Figure 10: Dataset 4 includes a combination of plant and animal sample images, without some of the conflicts uncovered when assessing previous datasets. The results show an overall improvement with inferences used to construct the subsequent, final dataset.

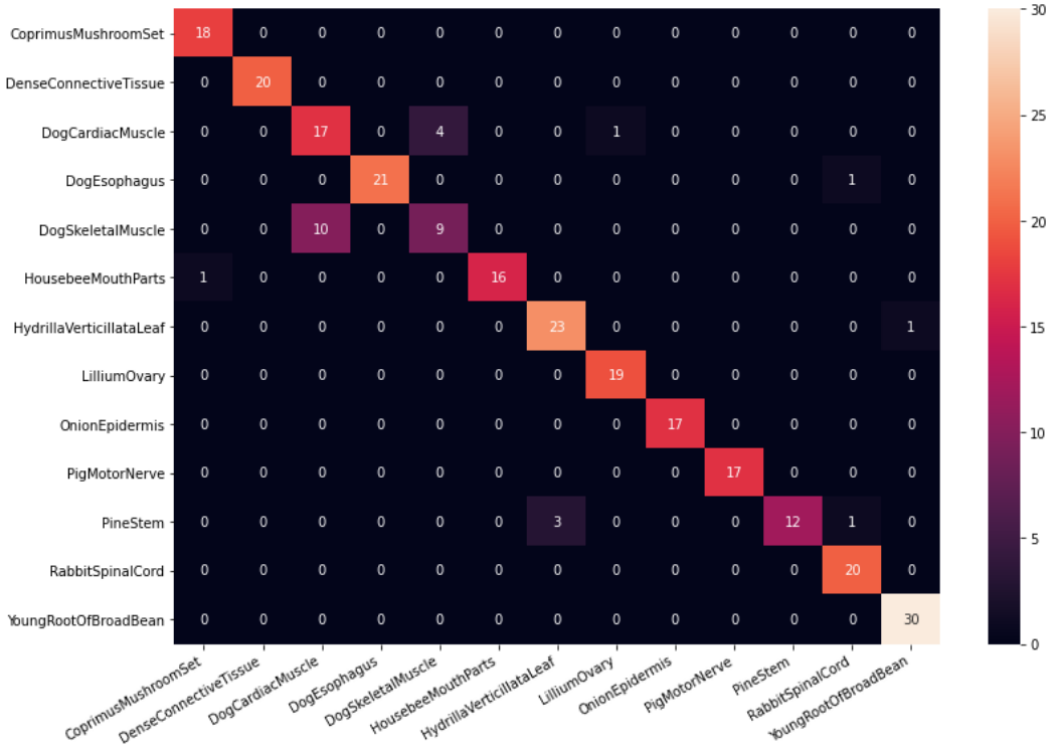

Figure 11: The confusion matrix reveals a greater degree of false identifications when using Dataset 4. The confusion is between "Dog Cardiac Muscle" and "Dog Skeletal Muscle"; and between plant sample images of "Hydrilla Verticillata Leaf" and "Pine Stem."

a maximum categorical accuracy of 97.09% and 92.95% for testing and validation subsets, respectively. The confusion matrix (Figure 9) reveals three false positives between "Sunflower Stem" and "Pumpkin Stem"; and three false positives and six false negatives when comparing "Sunflower Stem" with "Young Root of Broad Bean."

4.4 Dataset 4

Through analysis of datasets 1, 2 and 3, multiple leaf and stem samples were removed leaving just the "Pine Stem" and "Hydrilla Verticillata Leaf" samples (refer to Table 1 for details). The rest of the samples, except dog muscles, are a unique part of a plant or animal. Running this dataset through the model produced a peak classification accuracy of 96.45% and

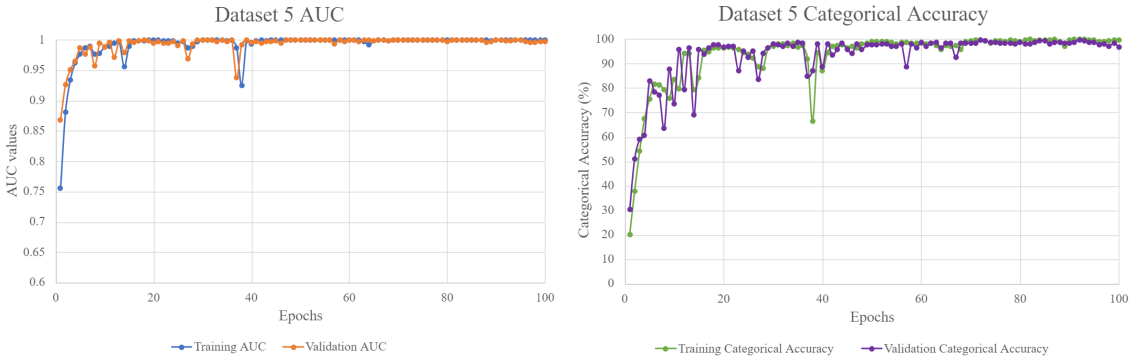

Figure 12: Dataset 5 is a subset of the original 19 sample images. The images that were identified in the confusion matrices for previous datasets are removed to optimise the dataset for the most effective identification.

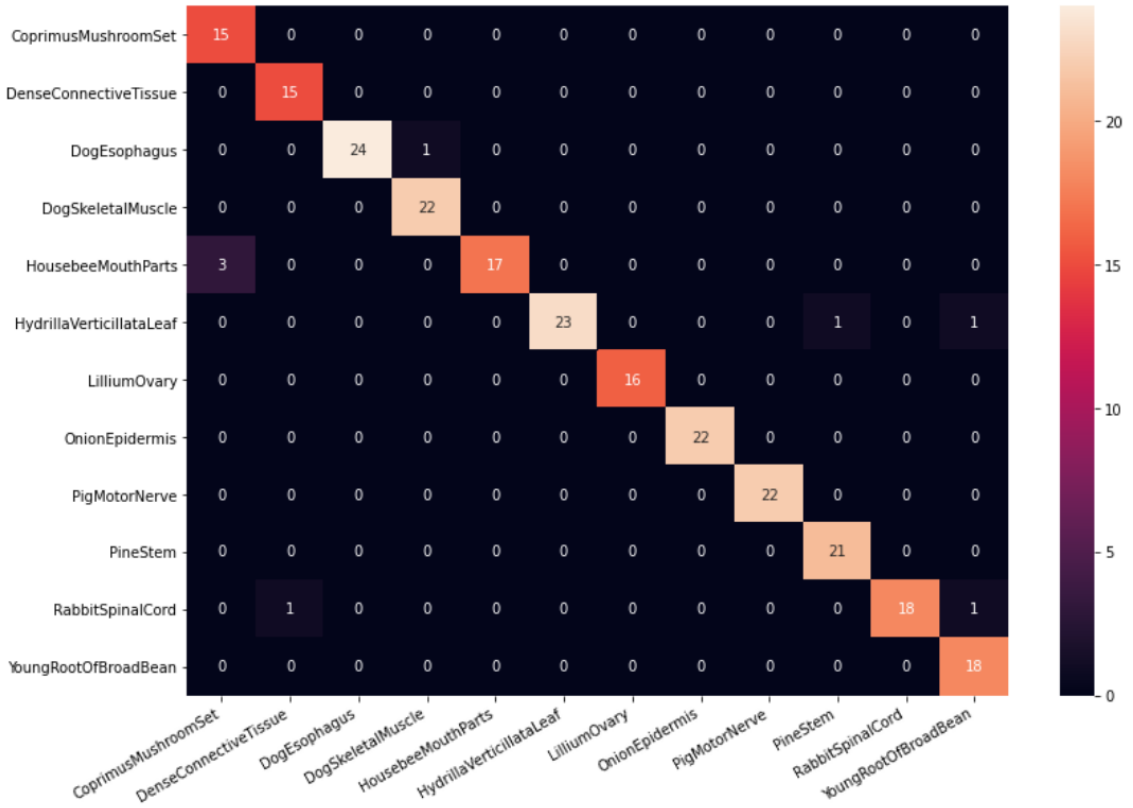

Figure 13: While Dataset 5 produces the best results, we still see conflicts with three false negatives between the "Coprimus Mushroom Set" and "House Bee Mouth Parts." The confusion matrix illustrates both the conflicts and the high degree of accurate classifications.

95.02% for the training and validation subsets, respectively. Plotting the AUC values for the test-train data against epochs (Figure 10) showed the subsets matching their fluctuations overall. The peak values for the training and testing subsets are 0.9996 and 0.999 respectively.

Plotting categorical accuracy as a percentage against epochs shows an upwards trend of fluctuating values with a steep decrease between epochs 94 and 97. Subsequently, an increase to greater than 91% for the last three epochs. This could be due to the random sample batches between epochs 94 and 97 containing multiple dog muscle samples. There was a peak categorical accuracy of 96.45% and 95.02% for the training and validation subsets, respectively.

The confusion matrix (Figure 11) reveals ten false negative and four false positive identifications between "Dog Cardiac Muscle" and "Dog Skeletal Muscle." There was also some confusion between the "Hydrilla Verticillata Leaf" and "Pine Stem" with three false negative results.

4.5 Dataset 5

Subsequently, we further refined our input dataset by removing "Dog Cardiac Muscle" images to produce our final dataset, i.e. dataset 5 (refer to Table 1 for details). This dataset achieved a peak categorical accuracy of 99.9% and 99.59% for the training and testing subsets, respectively. This confirms that dataset 5 produces the best results when applied to the CNN. Plotting the AUC values for the test-train data against epochs (Figure 12) showed an exceptionally efficient CNN with maximum AUC values of 1 for both subsets. At epochs 37 and 38, there is a brief drop in probability (of a correct classification) with values still above 0.9.

Plotting categorical accuracy as a percentage against epochs shows values for both test-train subsets to be over 90% for 70 epochs with peak categorical accuracy of 99.9% and 99.59% for training and testing subsets, respectively. At epoch 38, similar to the AUC plots, there is a sudden decrease in categorical accuracy equal to 66.49% for the training subset. The confusion matrix (Figure 13) revealed three false negative classifications between the "Coprimus Mushroom Set" and "House Bee Mouth Parts" sample images.

5 CONCLUSIONS

This research has demonstrated the biological datasets capabilities when applied to a CNN classification algorithm. With inferences drawn from classifying different subsets of the image set, we found a higher peak categorical accuracy of 99.9% and 99.59% when assessing dataset 5's training and validation subsets. The CNN consisting of two convolutional layers, with a SoftMax output has been shown to work efficiently and accurately. We also show evidence that the 3D printed OpenFlexure delta stage microscope is a viable replacement to standard microscopes for data collection (within the scope of this study). We reiterate our confidence in the benefits of using various open-source resources to develop an effective solution.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support of Dr Richard Bowman and his OpenFlexure team (University of Bath) for assistance during the construction and testing of the Openflexure microscope. We also wish to acknowledge the support from the EPSRC Research Council funding to QuantIC (EP/M01326X/1).

REFERENCES

[1] Organization, W. H., [*Malaria microscopy quality assurance manual-version 2*], World Health Organization (2016).

[2] "How Cancer Is Diagnosed - National Cancer Institute." https://www.cancer.gov/about-cancer/diagnosis-staging/diagnosis. (Accessed on 07/07/2021).

[3] Valdivia, R. H., Hromockyj, A. E., Monack, D., Ramakrishnan, L., and Falkow, S., "Applications for green fluorescent protein (GFP) in the study of hostpathogen interactions," *Gene* **173**(1) (1996).

[4] Wu, V. M.-T. and Uskoković, V., "Fruit fly as a model organism for blood-brain barrier penetration and infectious disease in the nanomedical niche," *Journal of Bionic Engineering* **17**, 553–569 (2020).

[5] Tersigni, M. A., "Frozen human bone: a microscopic investigation," *Journal of Forensic Sciences* **52**(1), 16–20 (2007).

[6] Koch, S. L., Liebowitz, C., Shriver, M. D., and Jablonski, N. G., "Microscopical discrimination of human head hairs sharing a mitochondrial haplogroup," *Journal of Forensic Sciences* **66**(1), 56–71 (2021).

[7] Goddard, A. J., Hillman, A. R., and Bond, J. W., "High resolution imaging of latent fingerprints by localized corrosion on brass surfaces," *Journal of Forensic Sciences* **55**(1), 58–65 (2010).

[8] Westlake, S. and Bugmann, G., *Computer Vision Detection of Malaria with a Convolutional Neural Network*, Master's thesis (2018).

[9] Veeraraghavan, R. and Silverstein, J., "Microscopes in education: Unlocking unseen worlds and undreamed-of futures," *Microscopy Today* **29**(2), 48–51 (2021).

[10] Mitiku, K., Mengistu, G., and Gelaw, B., "The reliability of blood film examination for malaria at the peripheral health unit," *Ethiopian Journal of Health Development* **17**, 197–204 (2004).

[11] Rottenfusser, R., Wilson, E. E., and Davidson, M. W., "Microscope Cleaning and Maintenance – Education in Microscopy and Digital Imaging (Zeiss Website Series, http://zeiss-campus.magnet.fsu.edu/articles/basics/care.html)."

[12] Khan, A. I., Shah, J. L., and Bhat, M. M., "Coronet: A deep neural network for detection and diagnosis of covid-19 from chest x-ray images," *Computer Methods and Programs in Biomedicine* **196**, 105581 (2020).

[13] Manogaran, G., Varatharajan, R., and Priyan, M., "Hybrid recommendation system for heart disease diagnosis based on multiple kernel learning with adaptive neuro-fuzzy inference system," *Multimedia tools and applications* **77**(4), 4379–4399 (2018).

[14] Alakwaa, W., Nassef, M., and Badr, A., "Lung cancer detection and classification with 3D convolutional neural network (3D-CNN)," *Lung Cancer* **8**(8), 409 (2017).

[15] "Openflexure delta stage." https://openflexure.org/projects/deltastage/ (2021). (Accessed on 07/09/2021).

[16] Archibald, R., Westlake, S., Gibson, G., and Kallepalli, A., "OPEN-BIOset: A dataset of microscope images collected using the OpenFlexure Delta Stage Microscope," (2021).

[17] Collins, J. T., Knapper, J., Stirling, J., Mduda, J., Mkindi, C., Mayagaya, V., Mwakajinga, G. A., Nyakyi, P. T., Sanga, V. L., Carbery, D., White, L., Dale, S., Jieh Lim, Z., Baumberg, J. J., Cicuta, P., McDermott, S., Vodenicharski, B., and Bowman, R., "Robotic microscopy for everyone: the OpenFlexure microscope," *Biomedical Optics Express* **11**(5), 2447–2460 (2020).

[18] Sharkey, J. P., Foo, D. C., Kabla, A., Baumberg, J. J., and Bowman, R. W., "A one-piece 3D printed flexure translation stage for open-source microscopy," *Review of Scientific Instruments* **87**(2), 025104 (1–7) (2016).

[19] Patton, B. R., Burke, D., Owald, D., Gould, T. J., Bewersdorf, J., and Booth, M. J., "Three-dimensional STED microscopy of aberrating tissue using dual adaptive optics," *Optics Express* **24**(8), 8862–8876 (2016).

[20] Vallejo Ramirez, P. P., Zammit, J., Vanderpoorten, O., Riche, F., Blé, F.-X., Zhou, X.-H., Spiridon, B., Valentine, C., Spasov, S. E., Oluwasanya, P. W., Goodfellow, G., Fantham, M. J., Siddiqui, O., Alimagham, F., Robbins, M., Stretton, A., Simatos, D., Hadeler, O., Rees, E. J., Ströhl, F., Laine, R. F., and Kaminski, C. F., "OptiJ: Open-source optical projection tomography of large organ samples," *Scientific Reports* **9**(1), 1–9 (2019).

[21] Rosenblatt, F., "The perceptron: A probabilistic model for information storage and organization in the brain," (1958).

[22] Fukushima, K., "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics* **36**(4), 193–202 (1980).

[23] LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P., "Gradient-based learning applied to document recognition," in [*Proceedings of the IEEE*], **86**(11), 2278–2323, IEEE (1998).

[24] Google, "Google Brain (https://research.google/teams/brain/)," (2021).

[25] Krizhevsky, A., Sutskever, I., and Hinton, G. E., "ImageNet classification with Deep convolutional neural networks," *Communications of the ACM* **60**(6), 84–90 (2017).

[26] Xu, B., Wang, N., Chen, T., and Li, M., "Empirical Evaluation of Rectified Activations in Convolutional Network," *arXiv* (2015).

[27] Simonyan, K. and Zisserman, A., "Very deep convolutional networks for large-scale image recognition," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings* , 1–14 (2015).

[28] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., and Chen, T., "Recent advances in convolutional neural networks," *Pattern Recognition* **77**, 354–377 (2018).

[29] Pise, N. and Kulkarni, P., "Algorithm selection for classification problems," *Proceedings of 2016 SAI Computing Conference, SAI 2016* , 203–211 (2016).

[30] Goodfellow, I., Bengio, Y., and Courville, A., [*Deep Learning*], MIT Press (2016). http://www.deeplearningbook.org.

[31] Sedjelmaci, H., Guenab, F., Senouci, S. M., Moustafa, H., Liu, J., and Han, S., "Cyber Security Based on Artificial Intelligence for Cyber-Physical Systems," (2020).

[32] Grigorescu, S., Trasnea, B., Cocias, T., and Macesanu, G., "A survey of deep learning techniques for autonomous driving," *Journal of Field Robotics* **37**(3), 362–386 (2020).

[33] Chinnery, T., Arifin, A., Tay, K. Y., Leung, A., Nichols, A. C., Palma, D. A., Mattonen, S. A., and Lang, P., "Utilizing Artificial Intelligence for Head and Neck Cancer Outcomes Prediction From Imaging," *Canadian Association of Radiologists Journal* **72**(1), 73–85 (2021).

[34] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., and Zheng, X., "TensorFlow: A system for large-scale machine learning," *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2016* , 265–283 (2016).

[35] Biswas, S. and Barma, S., "A large-scale optical microscopy image dataset of potato tuber for deep learning based plant cell assessment," *Scientific Data* **7**(371), 1–11 (2020).

[36] Krizhevsky, A. and Hinton, G., "Learning multiple layers of features from tiny images," (2009).

[37] Nwankpa, C. E., Ijomah, W., Gachagan, A., and Marshall, S., "Activation functions: Comparison of trends in practice and research for deep learning," in [*2nd International Conference on Computational Sciences and Technology*], 124 – 133 (2020).

[38] Kingma, D. P. and Ba, J. L., "Adam: A method for stochastic optimization," in [*3rd International Conference on Learning Representations, ICLR 2015*], Bengio, Y. and LeCun, Y., eds., 1–15 (2015).

[39] Fawcett, T., "An introduction to ROC analysis," *Pattern Recognition Letters* **27**(8), 861–874 (2006).