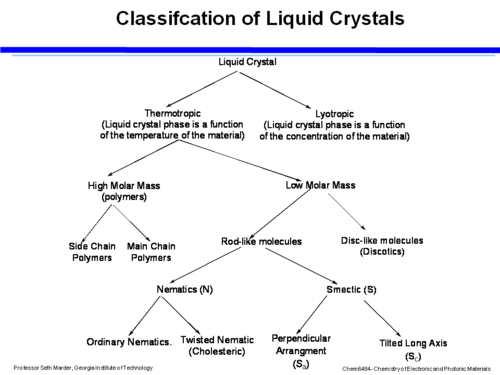
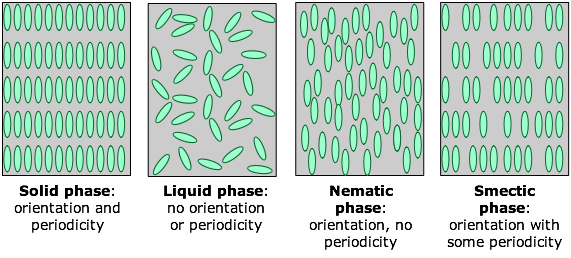
**Can liquid crystal phases be identified via machine learning?**

12/10/2020

The aim of this project is to test the usefulness of modern machine learning, in particular deep learning, algorithms in automating the identification of liquid crystal phases from their textures. A diagram of the different types and phases of liquid crystals is given below.





Liquid crystal texture image data is captured experimentally using polarized optical microscopy (POM). In the isotropic liquid phase with no molecular alignment (no director), light passes through without any change to its polarization axis. As the sample cools, it transitions into the liquid crystal phase at some critical temperature in which the molecules align along a director, leading to birefringent polarization of light passing through the sample. Therefore, when the sample is placed between crossed polarizers, incident light will only pass through when it is in a liquid crystal phase.

Since the data is in image format, a convolution neural network (CNN) will most likely be used in all cases of classification or regression in this project. Some potential applications:

* Phase classification of isotropic/nematic/cholesteric/smectic liquid crystals
* Cholesteric liquid crystal pitch regression
* Nematic liquid crystal temperature regression

**Data sources**

**Experimental**

Vance Williams

<https://www.instagram.com/vance.williams/>

<https://www.youtube.com/channel/UCB8qnCxJbdsuXpQ5RbLNy3Q>

<http://www.sfu.ca/chemistry/people/profiles/vancew/>

Media VR Lab YouTube channel

<https://www.youtube.com/channel/UCqb11FvjkkpD0V5h6p9c2Cg>

Nematic texture video

<https://www.youtube.com/watch?v=c4FuNSUHAPU>

**Simulations**

Nematic textures could potentially be simulated via Monte Carlo methods.

16/10/2020

**First model, Sequential 4 Phases**

**Collecting and processing the data**

We will first attempt to classify images from the following general phases, with an example image provided for each:

* Isotropic (random dark noise due to no birefringence)

A picture containing dark, front, standing, water

Description automatically generated

* Nematic

A picture containing building

Description automatically generated

* Smectic

A picture containing clothing, fabric, rug

Description automatically generated

* Cholesteric

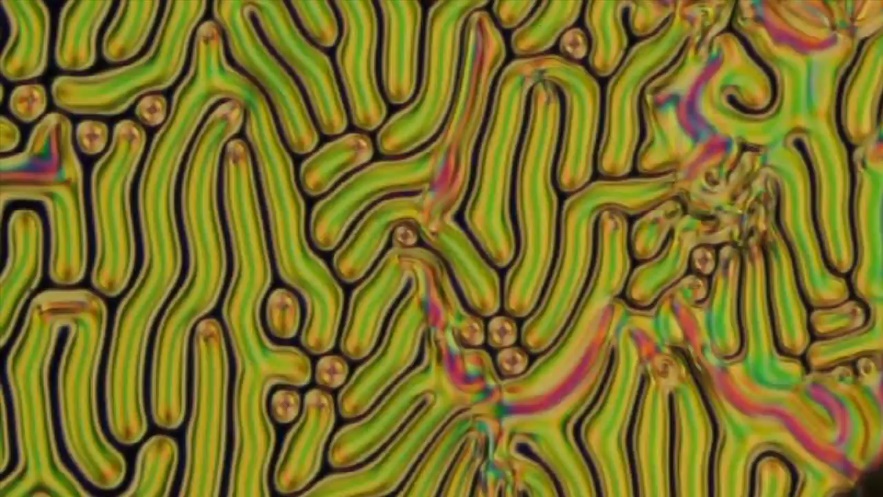
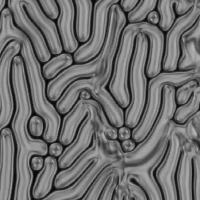
A picture containing honeycomb

Description automatically generated

We will use data from POM videos from Vance William’s YouTube and Instagram as well as some nematic videos from supervisor Ingo Dierking. Frames are extracted from the videos using the VLC Player software’s scene filter function. The isotropic phase is almost completely dark, so we generate random dark noise images for this class. Using this method, we obtained a total of 1742 texture images, not including isotropic images, and generated 500 isotropic images. We have written a script to process the images that does as follows:

1. Crops any excess pixels from the image in order to make it square, with the side length equal to maximum of width/height. This is to keep the proportions correct when resizing. Pixels are removed equally from either side.
2. Resizes the image to a given square dimension, *d*.
3. Optionally converts the image to grayscale.

Here is an example transformation of a cholesteric texture from 640x360 RGB to 200x200 grayscale:



18/10/2020

**Implementation of the models**

We chose an input image size of 200x200 and split the overall image data into a 7:2:1 ratio of training to validation to test sets. We also chose to convert to grayscale because colour is not important in determining LC phases, and it will reduce model size and training time. The pixel data is rescaled from 8 bit unsigned integer (0-255) to float range 0-1. The finalised distribution of data is as follows:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Isotropic | Nematic | Cholesteric | Smectic | Totals |
| Training | 350 | 475 | 372 | 375 | 1572 |
| Validation | 100 | 135 | 106 | 106 | 447 |
| Test | 50 | 67 | 53 | 53 | 223 |
| Totals | 500 | 677 | 531 | 534 | 2242 |

For the sequential models V1 and V2 we use the following settings.

Training specifications:

* Batch size: 32
* Loss function: categorical cross entropy
* Optimiser: Adam, initial learning rate = 0.001
* All training done using TensorFlow Keras with NVIDIA CUDA on an NVIDIA RTX 2060 GPU.

Keras ImageDataGenerator augmentations to improve regularisation:

* Vertical flip
* Horizontal flip
* 30 rotation range
* 0.1 height shift range
* 0.1 width shift range
* 0.2 zoom range

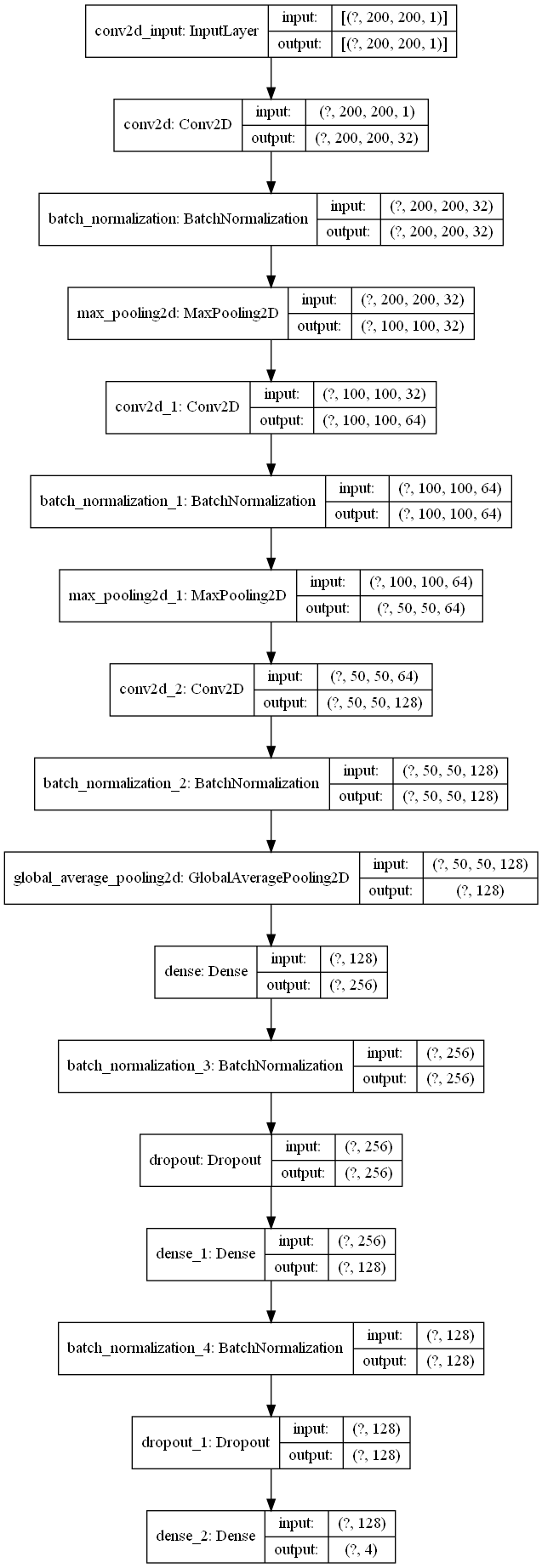
We apply early stopping based on validation set loss with patience of 30 epochs for improved regularisation.

We also reduce the learning rate during training if the validation loss does not improve within 10 epochs, with a reduction factor of 0.5 and minimum learning rate of 1e-5.

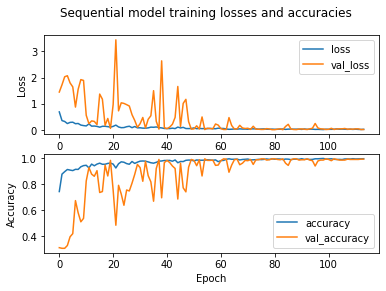
19/10/2020

**Sequential V1**

The first model, Sequential V1, has 3 convolution layers all with ReLU activations and batch normalisation, with max pooling used after the first 2 layers and global average pooling for the last. Each convolutional layer has a stride of 1, kernel size 3x3 and “same” padding. This is followed with 2 dense layers with ReLU activations and dropout for regularisation, and a final 4 unit dense layer for the output. The entire architecture with channel and unit numbers is displayed below.

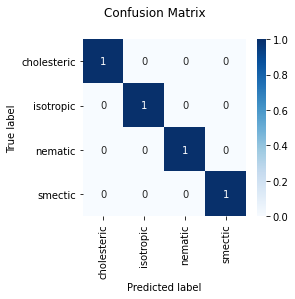


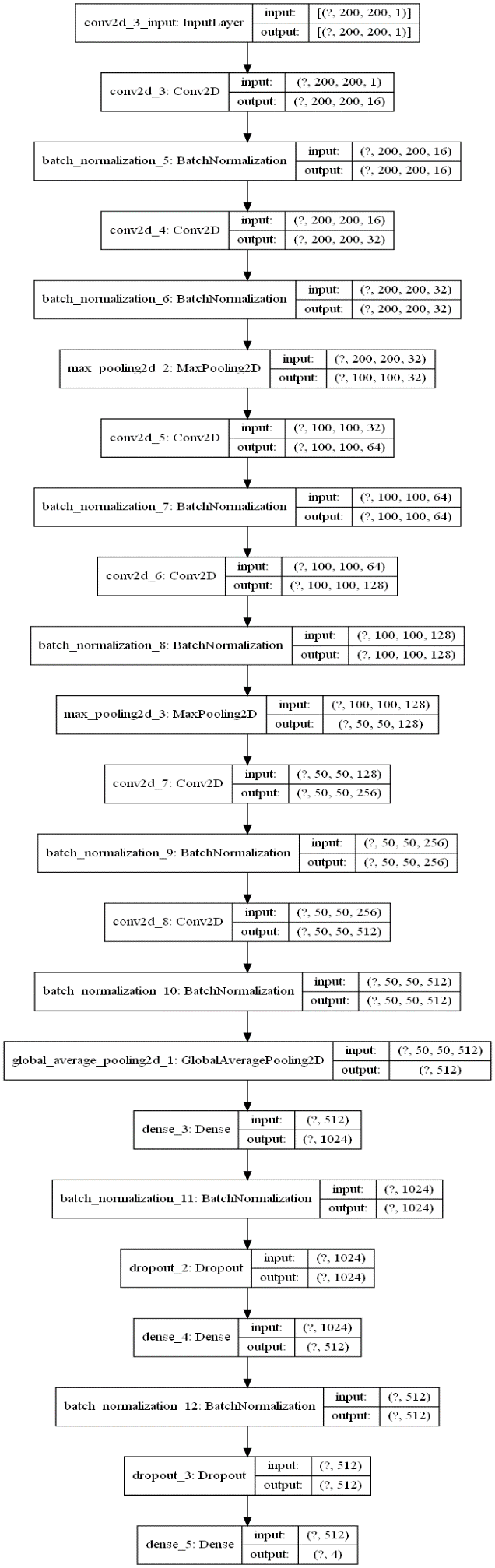
**Sequential V1**

After completion of the training cycle, the final validation set accuracy was 99.04%, and test set accuracy was 100.00%. This clearly demonstrates that CNNs will be extremely viable for the LC phase classification task, as this was a relatively small model. The graphs of loss and accuracy for both training and validation sets over the entire training period are displayed below.

The validation accuracy fluctuates greatly at the beginning before converging to near 100%.

The confusion matrix for the test set predictions is displayed below.





**Sequential V2**

The model has 100% accuracy on the test set leading to a fully diagonal confusion matrix.

20/10/2020

**Sequential V2**

The second sequential model is similar to V1, but it has 3 extra convolutional layers and more filters and dense units, for a much higher total model capacity. The full architecture is displayed to the left.

The model achieved a higher validation set accuracy of 99.76%, and again a test set accuracy of 100.00%. The accuracies for both models are potentially extremely high because image samples from the same videos were included in the training, validation and test sets, which could be a form of data leakage as images from the same video can be very similar to each other. This could mean the models do not actually generalise well. Next, we will try training on a dataset in which training images are from completely different videos to the validation and test set ones, in order to avoid any form of data leakage.

The V2 model loss and accuracy plot is displayed below.

Chart, histogram

Description automatically generated

8/11/2020

**Sequential V3**

For this model we will explicitly experiment with the effect of increasing the number of convolutional layers on model accuracy. We now have an expanded dataset with POM videos from the LC compounds:

* 5CB
* 8CB
* D5
* D6
* D7
* D8
* M5
* M6
* M7
* M9
* M10

We choose to try input image sizes of 256x256 and 128x128 (powers of 2 help with GPU memory). The new videos have large dimensions of 2048x1088 pixels. We will split each one into 6 sub images of size 682x544, before cropping to square and resizing as with the previous set of images. Also, when splitting the data into train, validation and test sets, we now avoid putting images from the same video in more than one set, in order to prevent the data leakage observed in models V1 and V2. The new distribution of data is as follows:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Isotropic | Nematic | Cholesteric | Smectic | Totals |
| Training | 1500 | 1691 | 1549 | 1689 | 6429 |
| Validation | 400 | 471 | 405 | 457 | 1733 |
| Test | 200 | 208 | 178 | 287 | 873 |
| Totals | 2100 | 2370 | 2132 | 2433 | 9035 |

We first train six individual models with convolutional layer number increasing from 1 to 6. For each model, the first layer has 32 filters, with the number of filters doubling with each successive convolutional layer. Each convolutional layer, like the first models, has ReLU activations, “same” padding, stride of 1 and kernel size 3x3. In addition, each convolutional layer is followed by batch normalization and then max pooling with pool size 2x2. The final convolutional layer is followed by global average pooling instead of max pooling. Each model has 2 dense layers, first with 256 units and then 128, both with dropout rate 0.5 for regularisation, as well as batch normalisation, followed by a 4 unit output layer. We use less dense units than V1 and V2 in order to prevent overfitting. We use the same training specifications as for V1 and V2, aside from an increased early stopping patience of 100 epochs. We also train the models a second time with only horizontal and vertical flip augmentations, for input image sizes of

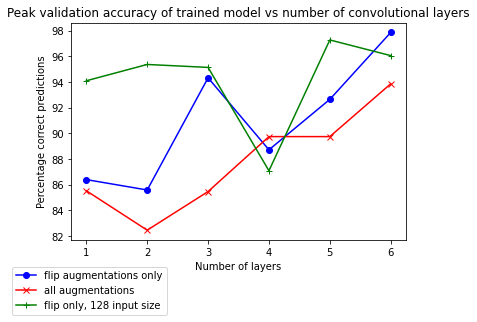
The results of the training are summarised in the tables below:

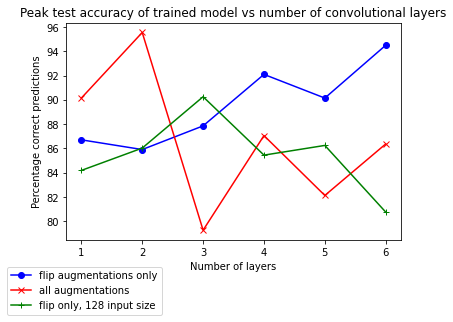
|  |  |  |  |
| --- | --- | --- | --- |
| **All image augmentations (same as models V1 and V2), input size 256x256** | | | |
| Number of convolutional layers | Validation accuracy (%) | Test accuracy (%) | Training time (s) |
| 1 | 85.53 | 90.15 | 6210 |
| 2 | 82.46 | 95.55 | 5520 |
| 3 | 85.47 | 79.27 | 5750 |
| 4 | 89.75 | 87.06 | 5474 |
| 5 | 89.75 | 82.13 | 6486 |
| 6 | 93.86 | 86.37 | 5290 |
| Mean | **87.80** | **86.76** | **5788** |

|  |  |  |  |
| --- | --- | --- | --- |
| **Vertical and horizontal flip augmentations only, input size 256x256** | | | |
| Number of convolutional layers | Validation accuracy (%) | Test accuracy (%) | Training time (s) |
| 1 | 86.40 | 86.71 | 3040 |
| 2 | 85.59 | 85.91 | 2180 |
| 3 | 94.33 | 87.86 | 3040 |
| 4 | 88.72 | 92.10 | 2289 |
| 5 | 92.65 | 90.15 | 3108 |
| 6 | 97.86 | 94.50 | 3990 |
| Mean | **90.92** | **89.54** | **2941** |

|  |  |  |  |
| --- | --- | --- | --- |
| **Vertical and horizontal flip augmentations only, input size 128x128** | | | |
| Number of convolutional layers | Validation accuracy (%) | Test accuracy (%) | Training time (s) |
| 1 | 94.10 | 84.19 | 2736 |
| 2 | 95.37 | 86.03 | 2470 |
| 3 | 95.14 | 90.26 | 3439 |
| 4 | 87.09 | 85.45 | 2261 |
| 5 | 97.28 | 86.25 | 3249 |
| 6 | 96.06 | 80.76 | 2926 |
| Mean | **94.17** | **85.49** | **2847** |

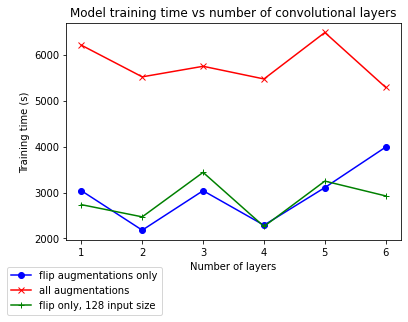
As can be seen from the mean accuracies, the models have a higher accuracy when trained using only flip augmentations. This is potentially due to the vastly expanded dataset requiring less augmentation for good regularisation. The 256x256 input models have similar validation and training accuracies in both cases, suggesting good model regularisation. The models with input size 128x128 have a higher mean validation accuracy than the 256x256 models, but the mean test set accuracy is much lower than the validation accuracy, suggesting that they are not well regularised. Since the training times do not differ by much for both input sizes with only flip augmentations, we will continue with future models using the 256x256 input size only. Plots of model accuracy on both validation and test sets vs number of convolutional layers are displayed below.





The general trend appears to be an increase in accuracy with the number of convolutional layers for the validation set, with not obvious trend for the test accuracies.

The mean training time using all augmentations is approximately double that with only flip augmentations, demonstrating that augmentation processing is a major computational expense. A plot of training time vs number of convolutional layers is displayed below.



There does not appear to be any correlation between training time and number of layers from this plot.

The architecture diagrams, training loss/accuracy graphs and test set prediction confusion matrices for each model with input size 256x256 are listed below.

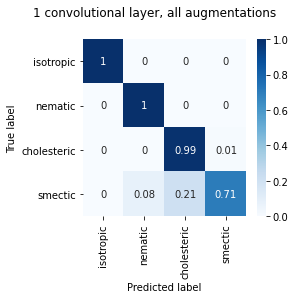
A picture containing graphical user interface

Description automatically generated**1 convolutional layer:**

All augmentations:

**A picture containing chart

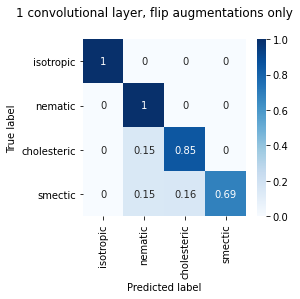
Description automatically generated**



Flip augmentations only:

Chart, histogram

Description automatically generated

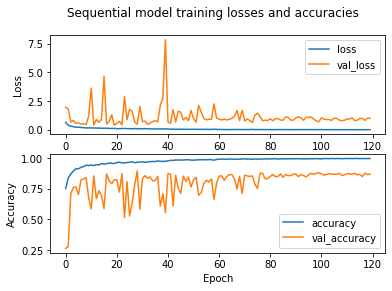


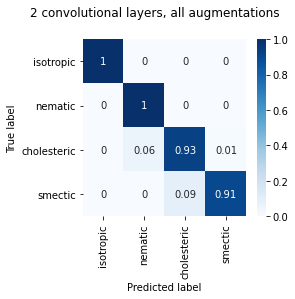
**2 convolutional layers:**

**A picture containing icon

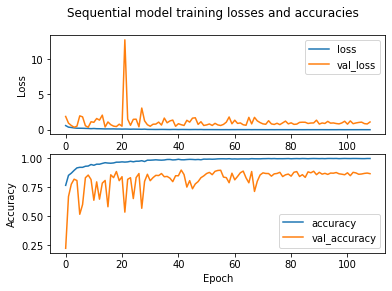
Description automatically generated**

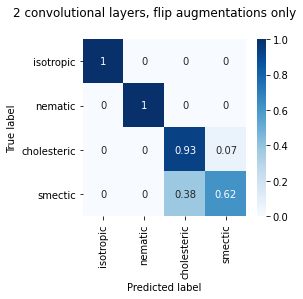
All augmentations:





Flip augmentations only:



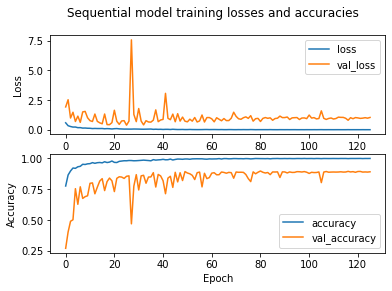


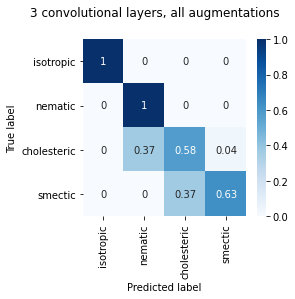
**3 convolutional layers:**

**A picture containing diagram

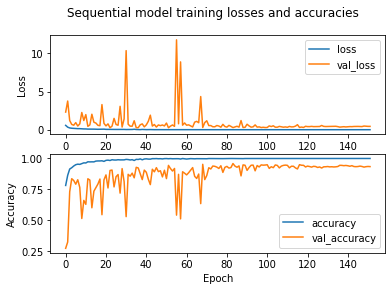
Description automatically generated**

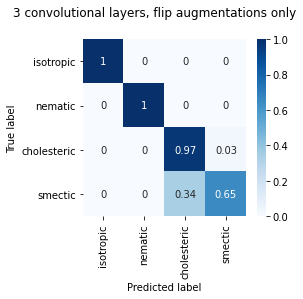
All augmentations:





Flip augmentations only:



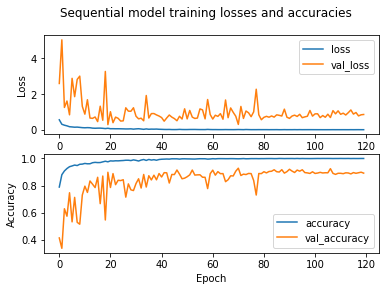


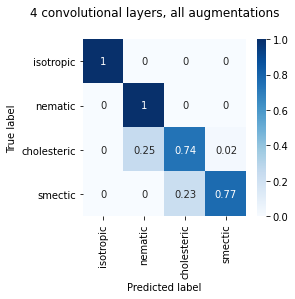
**4 convolutional layers:**

**A picture containing icon

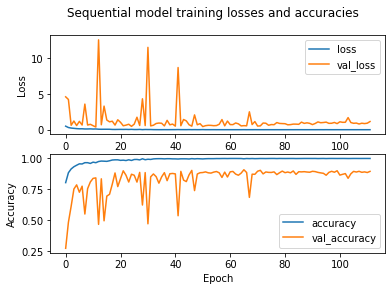
Description automatically generated**

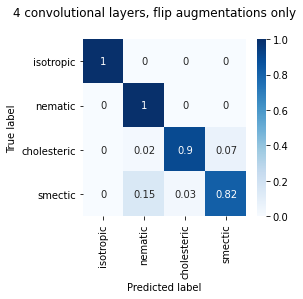
All augmentations:





Flip augmentations only:



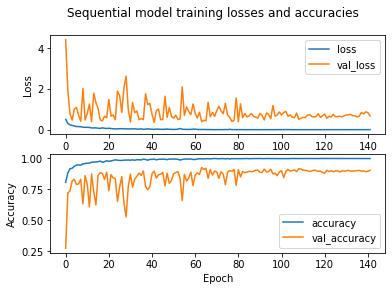


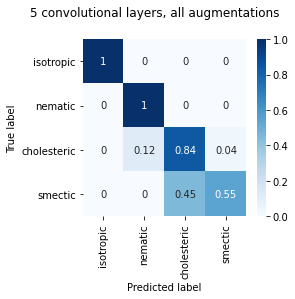
**5 convolutional layers:**

**Diagram

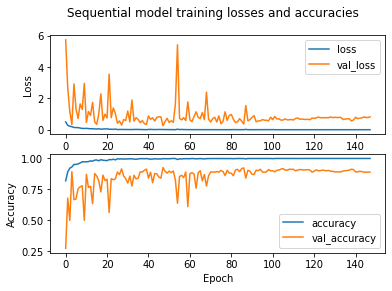
Description automatically generated**

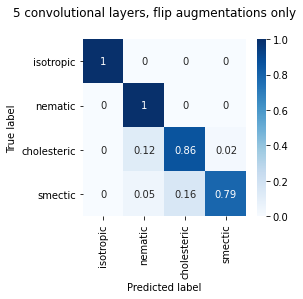
All augmentations:





Flip augmentations only:



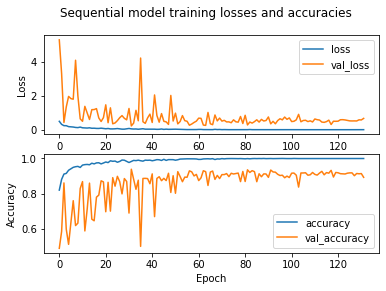


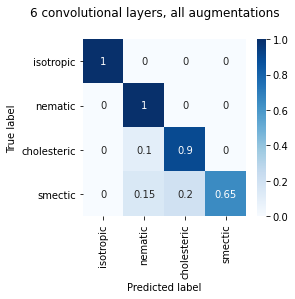
**6 convolutional layers:**

**A close up of a sign

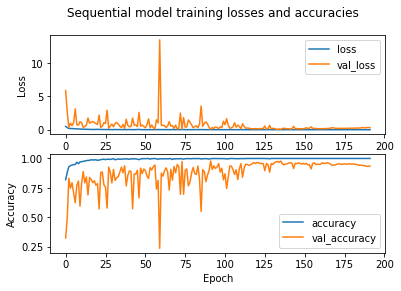
Description automatically generated**

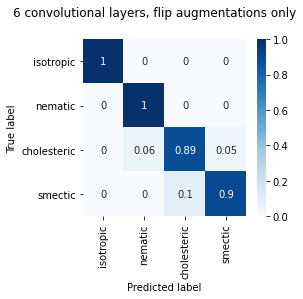
All augmentations:





Flip augmentations only:





It is clear from the confusion matrices that all models are highly accurate in identifying the isotropic and nematic phases whilst often confusing the cholesteric and smectic phases. This is potentially due to the similarities between certain cholesteric and smectic texture images.

17/11/2020

**Repeating the training**

We will now repeat all V3 training 2 more times and measure the accuracies in order to calculate a mean accuracy and uncertainty for each model, this time including the case with all augmentations and 128x128 image size. Uncertainty on the mean is taken as half the range of the values, due to the small sample size of 3.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **All augmentations, 256x256 input size, validation set accuracies** | | | | | | |
| Number of conv. layers | 1 | 2 | 3 | 4 | 5 | 6 |
| 1st run | 85.53 | 82.46 | 85.47 | 89.75 | 89.75 | 93.86 |
| 2nd run | 88.14 | 87.44 | 89.53 | 88.14 | 97.51 | 92.07 |
| 3rd run | 85.07 | 86.46 | 87.44 | 86.23 | 91.78 | 96.76 |
| Means | 86.25 | 85.45 | 87.48 | 88.04 | 93.01 | 94.23 |
| Uncertainty | 1.54 | 2.49 | 2.03 | 1.76 | 3.88 | 2.35 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **All augmentations, 256x256 input size, test set accuracies** | | | | | | |
| Number of conv. layers | 1 | 2 | 3 | 4 | 5 | 6 |
| 1st run | 90.15 | 95.55 | 79.27 | 87.06 | 82.13 | 86.37 |
| 2nd run | 91.78 | 81.25 | 78.36 | 83.22 | 81.83 | 91.20 |
| 3rd run | 90.28 | 93.98 | 78.47 | 84.72 | 82.41 | 87.27 |
| Means | 90.74 | 90.26 | 78.70 | 85.00 | 82.12 | 88.28 |
| Uncertainty | 0.81 | 7.15 | 0.45 | 1.92 | 0.29 | 2.42 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Flip augmentations, 256x256 input size, validation set accuracies** | | | | | | |
| Number of conv. layers | 1 | 2 | 3 | 4 | 5 | 6 |
| 1st run | 86.40 | 85.59 | 94.33 | 88.72 | 92.65 | 97.86 |
| 2nd run | 88.48 | 83.56 | 92.65 | 86.57 | 93.34 | 91.96 |
| 3rd run | 85.13 | 81.13 | 91.44 | 86.92 | 94.85 | 86.34 |
| Means | 86.67 | 83.43 | 92.81 | 87.40 | 93.61 | 92.05 |
| Uncertainty | 1.68 | 2.23 | 1.45 | 1.08 | 1.10 | 5.76 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Flip augmentations, 256x256 input size, test set accuracies** | | | | | | |
| Number of conv. layers | 1 | 2 | 3 | 4 | 5 | 6 |
| 1st run | 86.71 | 85.91 | 87.86 | 92.10 | 90.15 | 94.50 |
| 2nd run | 89.00 | 94.33 | 88.19 | 86.46 | 86.69 | 84.49 |
| 3rd run | 97.57 | 94.21 | 80.90 | 81.37 | 95.83 | 89.47 |
| Means | 91.09 | 91.48 | 85.65 | 86.64 | 90.89 | 89.49 |
| Uncertainty | 5.43 | 4.21 | 3.65 | 5.36 | 4.57 | 5.01 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **All augmentations, 128x128 input size, validation set accuracies** | | | | | | |
| Number of conv. layers | 1 | 2 | 3 | 4 | 5 | 6 |
| 1st run | 82.06 | 86.63 | 84.32 | 93.34 | 91.32 | 91.09 |
| 2nd run | 88.37 | 88.54 | 88.95 | 90.16 | 95.49 | 93.00 |
| 3rd run | 90.34 | 84.14 | 87.96 | 89.93 | 93.06 | 94.50 |
| Means | 86.92 | 86.44 | 87.08 | 91.14 | 93.29 | 92.86 |
| Uncertainty | 4.14 | 2.20 | 2.32 | 1.71 | 2.09 | 1.71 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **All augmentations, 128x128 input size, test set accuracies** | | | | | | |
| Number of conv. layers | 1 | 2 | 3 | 4 | 5 | 6 |
| 1st run | 83.56 | 86.69 | 90.39 | 80.9 | 76.39 | 74.42 |
| 2nd run | 88.43 | 82.64 | 75.69 | 81.02 | 79.75 | 76.85 |
| 3rd run | 90.39 | 93.87 | 89.24 | 85.3 | 78.47 | 89.12 |
| Means | 87.46 | 87.73 | 85.11 | 82.41 | 78.20 | 80.13 |
| Uncertainty | 3.42 | 5.62 | 7.35 | 2.20 | 1.68 | 7.35 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Flip augmentations, 128x128 input size, validation set accuracies** | | | | | | |
| Number of conv. layers | 1 | 2 | 3 | 4 | 5 | 6 |
| 1st run | 94.10 | 95.37 | 95.14 | 87.09 | 97.28 | 96.06 |
| 2nd run | 87.96 | 87.44 | 87.38 | 92.19 | 93.00 | 93.23 |
| 3rd run | 87.73 | 90.57 | 95.89 | 93.23 | 95.95 | 95.14 |
| Means | 89.93 | 91.13 | 92.80 | 90.84 | 95.41 | 94.81 |
| Uncertainty | 3.19 | 3.97 | 4.26 | 3.07 | 2.14 | 1.42 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Flip augmentations, 128x128 input size, test set accuracies** | | | | | | |
| Number of conv. layers | 1 | 2 | 3 | 4 | 5 | 6 |
| 1st run | 84.19 | 86.03 | 90.26 | 85.45 | 86.25 | 80.76 |
| 2nd run | 92.59 | 93.40 | 96.88 | 82.52 | 87.96 | 83.22 |
| 3rd run | 95.25 | 90.86 | 84.84 | 86.11 | 92.01 | 86.00 |
| Means | 90.68 | 90.10 | 90.66 | 84.69 | 88.74 | 83.33 |
| Uncertainty | 5.53 | 3.69 | 6.02 | 1.80 | 2.88 | 2.62 |

The means and uncertainties are displayed in the plots below.

Chart, line chart

Description automatically generatedChart, line chart

Description automatically generated

The overall trend for validation accuracy appears to be an increase in accuracy with number of layers, with the flip augmentations and 128x128 image size performing best. The test set appears to have a slight negative correlation, potentially due to poor model regularisation in the higher numbers of layers.

26/11/20

**Second model, smectic phase classification**

We now attempt to classify smectic textures from the following phase categories:

* Fluid smectic (smectic A and C)
* Hexatic (smectic I and F)
* Soft crystal (smectic X)

This could prove to be more challenging due to the greater similarity between the images from the categories. Image splitting, cropping and pre-processing is carried out in the same way as for the V3 models, using frame data from the same set of videos. We will first use images of size 256x256 and only flip augmentations. The data is distributed as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Fluid smectic | Hexatic | Soft crystal | Totals |
| Training | 1759 | 486 | 600 | 2845 |
| Validation | 372 | 90 | 144 | 606 |
| Test | 389 | 90 | 96 | 575 |
| Totals | 2520 | 666 | 840 | 4026 |

The current dataset is highly imbalanced in the favour of the fluid smectic class due to a limited number of hexatic and soft crystal videos. We first train some sequential models similar to the ones for the 4 phase classification, but with reduced overall capacity to prevent overfitting on the small dataset. After trying 1-6 convolutional layers, the common result is mode collapse into the fluid smectic class due to the imbalance, resulting in test set confusion matrices like this:

Chart

Description automatically generated

We will instead try a model based on Google’s “Inception” architecture, in which multiple convolutional layers with different kernel sizes are placed in parallel in order to extract features of different sizes, before being concatenated together. We will refer to a structure of this type as an inception block. We use flip augmentations only and 256x256 image size and the same training configurations as for V3. Representations of models with increasing numbers of inception blocks from 1 to 3 are displayed below, along with their accuracies and test set confusion matrices.

**Table

Description automatically generated1 inception block**:

Validation accuracy: 82.29%

Test accuracy: 72.24%

Chart

Description automatically generated

**Table

Description automatically generated2 inception blocks:**

Validation accuracy: 78.13%

Test accuracy: 83.27%

Chart, waterfall chart

Description automatically generated

**A close up of text on a black background

Description automatically generated3 inception blocks:**

Validation accuracy: 81.01%

Test accuracy: 85.48%

Chart

Description automatically generated

As can be seen from the confusion matrices, the models are relatively accurate at identifying the fluid smectic class, most likely due to the large imbalance in favour of fluid smectic. The model with 3 inception blocks performed best overall in terms of accuracy. We may need more data for the hexatic and soft crystal classes to improve these models further.

10/12/20

**Smectic A/C binary classifier**

We shall now investigate the binary classification of the fluid smectic phases A and C. These phases’ structures do not differ by much, resulting in similar textures. The distribution of the data set is as follows:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Smectic A | Smectic C | Totals |
| Training | 719 | 1067 | 1786 |
| Validation | 174 | 183 | 357 |
| Test | 204 | 198 | 402 |
| Totals | 1097 | 1448 | 2545 |

We will use 256x256 image size and flip augmentations only. We will test the same inception models as the previous smectic classification (aside from having 1 unit in the final dense layer as it is a binary classification task), as well the V3 sequential models but with reduced channels in each layer to reduce overfitting. The new number of channels for each model in each layer, in order of convolutional layers is:

* 1 conv layer: 32
* 2 conv layers: 16, 32
* 3 conv layers: 8, 16, 32
* 4 conv layers: 4, 8, 16, 32
* 5 conv layers: 2, 4, 8, 16, 32
* 6 conv layers: 1, 2, 4, 8, 16, 32

The only change to the usual training specifications is the loss function, which will now be the binary cross entropy instead of categorical cross entropy. Each model is trained 3 times, the results are displayed in the tables below. The uncertainty is calculated as half the range as before.

|  |  |  |  |
| --- | --- | --- | --- |
| **Inception models, validation set accuracies** | | | |
| No. of inception blocks | 1 | 2 | 3 |
| 1st run | 92.05 | 98.3 | 98.58 |
| 2nd run | 98.01 | 92.61 | 100 |
| 3rd run | 100 | 100 | 91.48 |
| Mean | 96.69 | 96.97 | 96.69 |
| Uncertainty | 3.98 | 3.69 | 4.26 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Inception models, test set accuracies** | | | |
| No. of inception blocks | 1 | 2 | 3 |
| 1st run | 95.57 | 96.35 | 96.61 |
| 2nd run | 95.57 | 94.53 | 94.27 |
| 3rd run | 95.83 | 99.22 | 93.49 |
| Mean | 95.66 | 96.70 | 94.79 |
| Uncertainty | 0.13 | 2.34 | 1.56 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sequential models, validation set accuracies** | | | | | | |
| No. of conv. layers | 1 | 2 | 3 | 4 | 5 | 6 |
| 1st run | 94.03 | 92.61 | 86.08 | 99.15 | 98.58 | 94.03 |
| 2nd run | 91.48 | 88.07 | 94.03 | 97.44 | 96.59 | 95.45 |
| 3rd run | 90.06 | 96.31 | 95.45 | 98.3 | 92.9 | 98.86 |
| Mean | 91.86 | 92.33 | 91.86 | 98.3 | 96.02 | 96.12 |
| Uncertainty | 1.99 | 4.12 | 4.69 | 0.85 | 2.84 | 2.41 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sequential models, test set accuracies** | | | | | | |
| No. of conv. layers | 1 | 2 | 3 | 4 | 5 | 6 |
| 1st run | 98.18 | 96.61 | 97.14 | 96.09 | 91.41 | 96.09 |
| 2nd run | 96.35 | 97.14 | 94.53 | 97.4 | 96.09 | 93.49 |
| 3rd run | 85.42 | 97.14 | 86.46 | 97.66 | 97.14 | 90.89 |
| Mean | 93.32 | 96.96 | 92.71 | 97.05 | 94.88 | 93.49 |
| Uncertainty | 6.38 | 0.26 | 5.34 | 0.78 | 2.86 | 2.6 |

The means and uncertainties are plotted below.

Chart, line chart

Description automatically generatedChart, line chart

Description automatically generated

All of the models performed well on this task, with none having an accuracy below 90%. The inception models performed best overall compared the sequential ones, but the single model with the highest mean accuracy in both validation and test sets was the 4 layer sequential one.

15/12/20

The confusion matrices below are for the saved models with the highest in (a) and lowest in (b) test set accuracies. The highest accuracy was a 2 block inception model at 99.22%, and the lowest was a 1 convolutional layer sequential model with 85.42%. The inception model’s accuracy is extremely high, potentially due to a “lucky” fluctuation. The sequential model misidentifies 35% of smectic A phases as smectic C, possibly due to a slight imbalance in the dataset in favour of smectic C.

**Chart, waterfall chart

Description automatically generated**

05/01/21

**Future work**

Some ways to extend the project next semester are:

* Expanding the dataset, especially the hexatic smectic and soft crystal phases, in order to train models that can classify a greater a number of phases with high accuracy.
* Investigate new types of state-of-the-art CNN architectures.
* Attempt to implement transformer networks, a different type of deep learning algorithm that has only recently begun to show high potential in image classification.
* If we can obtain the correct labelled data, regression models for properties of liquid crystals such as cholesteric pitch length could be implemented.

08/02/21

**Timeline for second semester**

1. Improve dataset
2. Consider more systematic uncertainty evaluation
3. Consider metrics for unbalanced datasets
4. Consider methods to tackle class imbalance
5. Implement and investigate transformer networks
6. Investigate further state of the art CNNs
7. Ambitious goal: accurate classifier of many phases

**Minimum uncertainty**

Going forward, a minimum percentage uncertainty will be assigned to model test accuracy predictions, as the square root of the total number of test set examples divided by itself.

**Transformer network implementation**

We have attempted to implement a transformer network following the paper “AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE”, however early testing has proved unsuccessful so far.

19/02/21

**Updated smectic 3 phase dataset**

We have now constructed an updated smectic dataset with additional samples in fluid and hexatic phases. The distribution of the new set is given below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Fluid smectic | Hexatic | Soft crystal | Totals |
| Training | 2047 | 1338 | 600 | 3985 |
| Validation | 576 | 474 | 144 | 1194 |
| Test | 581 | 414 | 96 | 1091 |
| Totals | 3204 | 2226 | 840 | 6270 |

Since there were no new samples in the soft crystal class, it is now substantially underrepresented.

26/02/21

**Inception models for the updated set**

We now train the inception models again with the updated smectic set. This time, we perform 12 training runs. The model setups and training configurations are the same as the previous attempt at the general smectic task, aside from an increased batch size of 64. With the larger sample size of 12, the uncertainty is now taken as the standard deviation. The results are presented below.

|  |  |  |  |
| --- | --- | --- | --- |
| **Inception models, validation set accuracies** | | | |
| Training run | Number of blocks | | |
|  | 1 | 2 | 3 |
| 1 | 69.1 | 62.24 | 69.44 |
| 2 | 70.66 | 61.37 | 79.77 |
| 3 | 64.41 | 59.55 | 71.88 |
| 4 | 71.61 | 54.6 | 67.19 |
| 5 | 70.92 | 66.06 | 50.52 |
| 6 | 55.9 | 71.01 | 58.33 |
| 7 | 59.98 | 60.68 | 58.68 |
| 8 | 68.4 | 64.58 | 66.06 |
| 9 | 57.29 | 53.99 | 48 |
| 10 | 65.62 | 68.92 | 65.54 |
| 11 | 64.41 | 55.3 | 57.03 |
| 12 | 72.48 | 61.81 | 60.85 |
| Mean | 65.9 | 61.68 | 62.77 |
| Uncertainty | 5.43 | 5.19 | 8.64 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Inception models, test set accuracies** | | | |
| Training run | Number of blocks | | |
|  | 1 | 2 | 3 |
| 1 | 75.83 | 62.87 | 62.68 |
| 2 | 65.53 | 66.82 | 81.34 |
| 3 | 69.94 | 83 | 85.48 |
| 4 | 74.08 | 72.52 | 84.19 |
| 5 | 71.51 | 87.13 | 47.79 |
| 6 | 62.04 | 69.85 | 64.71 |
| 7 | 68.93 | 76.93 | 65.9 |
| 8 | 76.84 | 84.93 | 87.59 |
| 9 | 71.51 | 68.57 | 53.4 |
| 10 | 66.64 | 77.3 | 76.19 |
| 11 | 60.94 | 68.38 | 57.17 |
| 12 | 67.56 | 72.06 | 83.36 |
| Mean | 69.28 | 74.2 | 70.82 |
| Uncertainty | 4.83 | 7.37 | 13.29 |

The means and uncertainties are displayed in the plot below.

Chart, box and whisker chart

Description automatically generated

The test set accuracy is noticeably higher than validation. This could imply that the test set is more similar to the training set than the validation set is. There is no notable improvement over the previous attempt at this task with the smaller dataset. We see large standard deviations suggesting low training stability.

01/03/21

**Focal Loss**

In an attempt to tackle the class imbalance issue, we retrain the best inception model from the updated smectic set using the focal loss, a type of loss function introduced by Lin et al in 2017 in the paper “Focal Loss for Dense Object Detection”. It is a modification of the standard cross-entropy that recues the loss for correctly classified examples, with a greater reduction for greater confidence in the prediction. This can help when training on an imbalanced set by ensuring the smaller classes are not “ignored” by the model. The focal loss function is

Alpha and gamma are the modulating and focussing hyperparameters respectively. We use the recommended default values of 0.25 for alpha and 2.0 for gamma, and train the 2 block inception model 12 times, with all other configurations the same as before.

|  |  |  |  |
| --- | --- | --- | --- |
| **Inception models with focal loss, validation set accuracies** | | | |
| Training run | Number of blocks | | |
|  | 1 | 2 | 3 |
| 1 | 66.06 | 57.73 | 48.35 |
| 2 | 57.55 | 76.48 | 54.08 |
| 3 | 63.02 | 48.26 | 70.23 |
| 4 | 66.58 | 53.47 | 48.09 |
| 5 | 65.97 | 65.02 | 69.27 |
| 6 | 61.46 | 56.86 | 48.18 |
| 7 | 61.63 | 58.25 | 39.93 |
| 8 | 59.64 | 71.01 | 61.28 |
| 9 | 69.62 | 55.82 | 62.93 |
| 10 | 56.42 | 48.44 | 58.85 |
| 11 | 60.5 | 64.06 | 63.54 |
| 12 | 63.11 | 53.04 | 67.1 |
| Mean | 62.63 | 59.04 | 57.65 |
| Uncertainty | 6.6 | 14.11 | 15.15 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Inception models with focal loss, test set accuracies** | | | |
| Training run | Number of blocks | | |
|  | 1 | 2 | 3 |
| 1 | 68.66 | 70.5 | 53.31 |
| 2 | 73.16 | 90.44 | 56.62 |
| 3 | 82.9 | 53.22 | 81.71 |
| 4 | 69.12 | 74.08 | 53.4 |
| 5 | 70.86 | 69.76 | 59.1 |
| 6 | 63.14 | 61.4 | 53.22 |
| 7 | 70.68 | 72.33 | 37.96 |
| 8 | 76.01 | 78.68 | 70.31 |
| 9 | 60.2 | 68.57 | 57.08 |
| 10 | 66.27 | 43.11 | 56.43 |
| 11 | 65.44 | 89.43 | 78.68 |
| 12 | 71.51 | 49.72 | 80.88 |
| Mean | 69.83 | 68.44 | 61.56 |
| Uncertainty | 11.35 | 23.67 | 21.88 |

The plot of the mean accuracies is displayed on the next page.

Chart, box and whisker chart

Description automatically generated

Overall the models with focal loss perform worse than the ones without, and have much larger deviations in accuracies. This suggests that focal loss is not suitable for this task. The best single trained models, however, did perform slightly better than the normal loss best models.

The mean confusion matrices (averaged over all training repeats) are displayed below.

Chart, waterfall chart

Description automatically generatedChart, bar chart, waterfall chart

Description automatically generatedChart, bar chart, waterfall chart

Description automatically generatedChart, waterfall chart

Description automatically generatedChart, bar chart, waterfall chart

Description automatically generatedChart, waterfall chart

Description automatically generated

05/03/21

**More new data sets**

We have created and split new datasets using all available images. These include (with a label in brackets):

* Cholesteric, smectic (ChSm2)
* Cholesteric, fluid smectic, hexatic smectic (ChSm3)
* Cholesteric, fluid smectic, hexatic smectic, soft crystal (ChSm4)
* Cholesteric, smectic A, C, F, I (ChSm5)
* Cholesteric, smectic A, C, F, I, soft crystal (ChSm6)
* Smectic A, C (smecticAC)
* Smectic I, F (smecticIF)

All have been split into training, validation and test with a 70:15:15 ratio.

**Batch size and learning rate on ChSm4 investigation**

We train the three layer sequential model, with 8 starting channels (doubling with each layer), for batch sizes of 16, 32 and 64, and learning rates of 1e-5, 1e-4, 1e-3, 1e-2, 1e-1. An early stopping patience of 30 is used, and the models are trained 5 times. The mean validation and test accuracies over the 5 runs for each configuration are presented in the tables below.

|  |  |  |  |
| --- | --- | --- | --- |
| **Validation mean accuracies** | | | |
| Learning rate | Batch size | | |
|  | 16 | 32 | 64 |
| 1.00E-05 | 64.85 | 62.65 | 64.49 |
| 1.00E-04 | 74.74 | 74.19 | 70.62 |
| 1.00E-03 | 72.24 | 68.96 | 71.76 |
| 1.00E-02 | 67.45 | 68.42 | 59.82 |
| 1.00E-01 | 53.52 | 58.16 | 50.77 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Test mean accuracies** | | | |
| Learning rate | Batch size | | |
|  | 16 | 32 | 64 |
| 1.00E-05 | 70.56 | 67.13 | 68.61 |
| 1.00E-04 | 78.91 | 77.1 | 77.88 |
| 1.00E-03 | 82.48 | 74.86 | 76.6 |
| 1.00E-02 | 74.86 | 76.96 | 68.27 |
| 1.00E-01 | 62.23 | 64.42 | 55.48 |

From these results, we choose batch size of 16 and learning rate of 1e-4 for future training.

19/03/21

**Training sequential and inception models on the new sets**

We now train 3 and 4 layer sequential models and 1, 2 and 3 block inception models on the ChSm3, smecticAC and smecticIF sets. Starting channels of 8, 16, and 32 are tested for each of the 3 and 4 layer sequential models, and 2, 4, and 8 for the inception models. We use a batch size of 16 and learning rate of 1e-4 in all cases, and perform 10 training runs for all models. All other training configurations and augmentations are as before.

**ChSm3, sequential**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Validation accuracies** | | | | | | |
| Run | Architecture (number of layers, starting channels) | | | | | |
|  | 3, 8 | 3, 16 | 3, 32 | 4, 8 | 4, 16 | 4, 32 |
| 1 | 84.01 | 78.31 | 88.42 | 84.38 | 88.24 | 82.54 |
| 2 | 78.95 | 83.82 | 83.36 | 79.5 | 84.74 | 83 |
| 3 | 80.51 | 83.09 | 86.03 | 80.79 | 81.8 | 85.57 |
| 4 | 81.53 | 76.84 | 84.74 | 70.59 | 87.13 | 88.69 |
| 5 | 85.66 | 89.34 | 84.1 | 81.53 | 84.28 | 82.26 |
| 6 | 83.46 | 83.27 | 81.71 | 82.9 | 83.55 | 86.76 |
| 7 | 78.03 | 89.43 | 87.68 | 81.99 | 84.1 | 82.63 |
| 8 | 88.79 | 86.21 | 87.13 | 71.97 | 83 | 87.68 |
| 9 | 84.1 | 84.19 | 86.4 | 87.87 | 84.47 | 80.61 |
| 10 | 80.15 | 78.58 | 80.7 | 86.03 | 84.83 | 80.33 |
| Mean | 82.52 | 83.31 | 85.03 | 80.75 | 84.61 | 84.01 |
| Uncertainty | 3.13 | 4.15 | 2.44 | 5.3 | 1.78 | 2.81 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Test accuracies** | | | | | | |
| Run | Architecture (number of layers, starting channels) | | | | | |
|  | 3, 8 | 3, 16 | 3, 32 | 4, 8 | 4, 16 | 4, 32 |
| 1 | 81.88 | 82.68 | 83.39 | 86.07 | 84.2 | 86.43 |
| 2 | 83.75 | 87.59 | 88.84 | 78.48 | 88.84 | 86.34 |
| 3 | 83.84 | 82.05 | 88.3 | 86.87 | 89.46 | 85.89 |
| 4 | 82.23 | 71.16 | 87.5 | 81.52 | 84.29 | 83.39 |
| 5 | 85.62 | 88.48 | 83.3 | 79.11 | 85.45 | 84.46 |
| 6 | 83.75 | 85.27 | 84.02 | 82.32 | 87.32 | 85.98 |
| 7 | 80.18 | 86.52 | 85.89 | 86.25 | 86.61 | 87.32 |
| 8 | 88.57 | 89.02 | 88.84 | 79.82 | 84.38 | 87.32 |
| 9 | 86.7 | 86.96 | 88.21 | 83.93 | 85.09 | 83.57 |
| 10 | 79.46 | 85.54 | 82.5 | 87.14 | 83.93 | 83.57 |
| Mean | 83.6 | 84.53 | 86.08 | 83.15 | 85.96 | 85.43 |
| Uncertainty | 2.69 | 4.95 | 2.42 | 3.18 | 1.91 | 1.47 |

Chart, box and whisker chart

Description automatically generated

Fairly good performance with 86.08% mean test accuracy for the best architecture. Test accuracies overall higher than validation.

Chart, bar chart, waterfall chart

Description automatically generated

The mean confusion matrix for the best shows that the inaccuracies are mostly due to confusion of the smectic classes.

**ChSm3, inception**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Validation accuracies** | | | | | | | | | |
| Run | Architecture (number of blocks, starting channels) | | | | | | | | |
|  | 1, 2 | 1, 4 | 1, 8 | 2, 2 | 2, 4 | 2, 8 | 3, 2 | 3, 4 | 3, 8 |
| 1 | 75.37 | 72.43 | 77.85 | 70.31 | 72.43 | 70.59 | 77.3 | 85.11 | 83.92 |
| 2 | 84.83 | 71.97 | 85.66 | 76.38 | 76.47 | 86.03 | 79.32 | 76.47 | 81.62 |
| 3 | 79.14 | 80.42 | 80.7 | 80.79 | 86.4 | 85.39 | 74.91 | 82.81 | 80.06 |
| 4 | 80.33 | 78.4 | 86.86 | 83.92 | 88.97 | 77.76 | 81.71 | 81.99 | 83 |
| 5 | 77.94 | 79.87 | 85.39 | 79.78 | 84.56 | 83.64 | 76.29 | 86.03 | 86.21 |
| 6 | 79.14 | 81.89 | 83 | 79.5 | 71.05 | 82.72 | 77.67 | 81.8 | 85.11 |
| 7 | 81.25 | 79.23 | 84.74 | 79.5 | 79.32 | 86.58 | 79.87 | 75.37 | 84.1 |
| 8 | 82.54 | 83.55 | 80.15 | 82.9 | 83 | 82.26 | 70.86 | 85.29 | 79.32 |
| 9 | 79.5 | 85.94 | 84.01 | 79.32 | 74.91 | 78.68 | 80.61 | 85.02 | 82.81 |
| 10 | 76.65 | 81.89 | 80.42 | 78.77 | 80.33 | 77.76 | 81.07 | 79.5 | 87.22 |
| Mean | 79.67 | 79.56 | 82.88 | 79.12 | 79.74 | 81.14 | 77.96 | 81.94 | 83.34 |
| Uncertainty | 2.63 | 4.22 | 2.8 | 3.55 | 5.72 | 4.71 | 3.15 | 3.57 | 2.4 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test accuracies** | | | | | | | | | |
| Run | Architecture (number of blocks, starting channels) | | | | | | | | |
|  | 1, 2 | 1, 4 | 1, 8 | 2, 2 | 2, 4 | 2, 8 | 3, 2 | 3, 4 | 3, 8 |
| 1 | 80.89 | 78.57 | 79.02 | 77.77 | 85.8 | 80.09 | 82.05 | 84.91 | 89.29 |
| 2 | 76.7 | 79.55 | 82.77 | 77.23 | 76.61 | 85 | 81.25 | 84.38 | 86.25 |
| 3 | 76.25 | 82.86 | 80.89 | 75.27 | 85 | 87.59 | 71.43 | 81.96 | 80.54 |
| 4 | 69.73 | 80.62 | 90.45 | 79.91 | 86.52 | 81.43 | 83.3 | 77.95 | 80.89 |
| 5 | 78.21 | 83.04 | 85.89 | 77.68 | 87.41 | 83.84 | 77.77 | 86.52 | 82.95 |
| 6 | 77.77 | 79.64 | 84.73 | 79.29 | 78.93 | 85.98 | 82.14 | 80.89 | 85.27 |
| 7 | 74.29 | 76.61 | 85.8 | 79.46 | 83.66 | 87.23 | 75.71 | 81.25 | 85.36 |
| 8 | 77.23 | 80.62 | 86.7 | 83.13 | 81.52 | 87.32 | 68.12 | 82.68 | 81.25 |
| 9 | 73.57 | 82.95 | 84.73 | 80.98 | 81.34 | 80.45 | 77.32 | 84.55 | 80.62 |
| 10 | 81.7 | 86.61 | 84.38 | 81.34 | 78.84 | 85.36 | 86.25 | 79.55 | 79.64 |
| Mean | 76.63 | 81.11 | 84.54 | 79.21 | 82.56 | 84.43 | 78.54 | 82.46 | 83.21 |
| Uncertainty | 3.33 | 2.69 | 3 | 2.18 | 3.49 | 2.72 | 5.33 | 2.52 | 3.01 |

**Chart, box and whisker chart

Description automatically generated**

The accuracies appear to increase with the number of starting channels, whereas increasing the number of inception blocks does not seem to provide any benefit. Overall performance was worse than for sequential.

Chart, bar chart, waterfall chart

Description automatically generated

This mean confusion matrix for the best inception model again shows substantial confusion between the smectic phases, in addition to a 16% misclassification of N\* as HSm.

**SmecticAC, sequential**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Validation accuracies** | | | | | | |
| Run | Architecture (number of layers, starting channels) | | | | | |
|  | 3, 8 | 3, 16 | 3, 32 | 4, 8 | 4, 16 | 4, 32 |
| 1 | 97.29 | 94.17 | 90.42 | 94.58 | 90.83 | 95.83 |
| 2 | 88.96 | 94.38 | 93.33 | 92.71 | 94.38 | 98.12 |
| 3 | 88.75 | 92.71 | 92.5 | 88.75 | 92.5 | 94.79 |
| 4 | 90.42 | 93.96 | 94.38 | 88.75 | 93.75 | 94.17 |
| 5 | 90.62 | 94.79 | 93.75 | 92.5 | 94.38 | 92.29 |
| 6 | 88.75 | 92.71 | 92.5 | 91.04 | 92.29 | 98.12 |
| 7 | 94.38 | 88.75 | 90.62 | 90.62 | 98.12 | 99.58 |
| 8 | 88.75 | 91.46 | 91.67 | 99.37 | 92.5 | 95.21 |
| 9 | 88.75 | 88.96 | 90.42 | 88.75 | 93.75 | 94.38 |
| 10 | 94.38 | 92.5 | 95.83 | 88.75 | 93.75 | 93.33 |
| Mean | 91.1 | 92.44 | 92.54 | 91.58 | 93.63 | 95.58 |
| Uncertainty | 2.95 | 2.03 | 1.73 | 3.24 | 1.83 | 2.22 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Test accuracies** | | | | | | |
| Run | Architecture (number of layers, starting channels) | | | | | |
|  | 3, 8 | 3, 16 | 3, 32 | 4, 8 | 4, 16 | 4, 32 |
| 1 | 96.88 | 95.51 | 99.8 | 100 | 98.83 | 91.8 |
| 2 | 92.77 | 99.61 | 97.07 | 98.05 | 97.27 | 97.66 |
| 3 | 98.63 | 98.24 | 98.24 | 100 | 97.66 | 90.04 |
| 4 | 96.68 | 97.66 | 99.8 | 96.48 | 91.99 | 98.05 |
| 5 | 99.41 | 97.07 | 90.82 | 97.66 | 99.02 | 94.14 |
| 6 | 100 | 99.61 | 99.8 | 100 | 97.07 | 89.26 |
| 7 | 99.22 | 99.61 | 99.8 | 98.24 | 93.75 | 97.27 |
| 8 | 97.27 | 90.62 | 92.38 | 99.22 | 99.8 | 92.19 |
| 9 | 99.61 | 99.8 | 98.83 | 99.22 | 98.63 | 94.73 |
| 10 | 100 | 98.83 | 98.24 | 98.05 | 98.83 | 99.61 |
| Mean | 98.05 | 97.66 | 97.48 | 98.69 | 97.29 | 94.47 |
| Uncertainty | 2.13 | 2.69 | 3.08 | 1.13 | 2.38 | 3.41 |

**Chart, box and whisker chart

Description automatically generated**

The sequential models perform extremely well on the smecticAC set, with 5 out 6 models above 97% mean test accuracy. The test accuracies are substantially higher than validation.

Chart, waterfall chart

Description automatically generated

**SmecticAC, inception**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Validation accuracies** | | | | | | | | | |
| Run | Architecture (number of blocks, starting channels) | | | | | | | | |
|  | 1, 2 | 1, 4 | 1, 8 | 2, 2 | 2, 4 | 2, 8 | 3, 2 | 3, 4 | 3, 8 |
| 1 | 96.04 | 95.21 | 98.75 | 89.38 | 99.58 | 96.67 | 96.67 | 97.5 | 100 |
| 2 | 100 | 96.04 | 100 | 98.75 | 94.38 | 98.54 | 90.83 | 91.46 | 96.25 |
| 3 | 92.5 | 98.12 | 93.12 | 97.92 | 96.67 | 97.5 | 92.5 | 100 | 98.75 |
| 4 | 91.25 | 94.38 | 95.63 | 94.38 | 94.38 | 99.37 | 97.29 | 100 | 94.38 |
| 5 | 99.79 | 98.33 | 94.38 | 94.79 | 99.58 | 92.5 | 93.96 | 94.58 | 100 |
| 6 | 88.75 | 97.92 | 95 | 91.04 | 100 | 99.58 | 96.04 | 99.37 | 94.79 |
| 7 | 93.75 | 88.75 | 97.08 | 93.96 | 100 | 93.96 | 94.38 | 95.63 | 97.08 |
| 8 | 88.75 | 98.96 | 94.58 | 95.42 | 96.25 | 98.54 | 98.96 | 94.58 | 97.71 |
| 9 | 93.96 | 94.17 | 98.54 | 100 | 96.25 | 100 | 98.12 | 96.04 | 93.33 |
| 10 | 92.29 | 99.58 | 93.12 | 91.67 | 98.33 | 99.79 | 93.96 | 93.96 | 98.75 |
| Mean | 93.71 | 96.15 | 96.02 | 94.73 | 97.54 | 97.65 | 95.27 | 96.31 | 97.1 |
| Uncertainty | 3.75 | 3.07 | 2.31 | 3.27 | 2.13 | 2.44 | 2.45 | 2.71 | 2.24 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test accuracies** | | | | | | | | | |
| Run | Architecture (number of blocks, starting channels) | | | | | | | | |
|  | 1, 2 | 1, 4 | 1, 8 | 2, 2 | 2, 4 | 2, 8 | 3, 2 | 3, 4 | 3, 8 |
| 1 | 98.05 | 99.8 | 90.23 | 100 | 87.5 | 98.63 | 91.02 | 97.85 | 99.22 |
| 2 | 98.63 | 99.02 | 95.31 | 99.02 | 99.61 | 98.63 | 93.75 | 99.61 | 98.83 |
| 3 | 100 | 99.41 | 95.12 | 99.8 | 95.7 | 93.95 | 99.8 | 99.41 | 97.27 |
| 4 | 100 | 100 | 98.05 | 100 | 98.83 | 89.65 | 99.41 | 93.95 | 99.8 |
| 5 | 85.74 | 98.05 | 98.44 | 99.8 | 93.75 | 95.31 | 99.22 | 99.61 | 94.53 |
| 6 | 99.02 | 91.99 | 99.8 | 96.88 | 84.96 | 99.02 | 99.8 | 85.55 | 99.61 |
| 7 | 99.02 | 98.44 | 92.38 | 97.27 | 99.8 | 93.16 | 97.85 | 99.61 | 95.12 |
| 8 | 100 | 96.88 | 99.41 | 99.8 | 95.12 | 86.33 | 97.07 | 99.8 | 99.22 |
| 9 | 97.66 | 98.05 | 89.65 | 99.8 | 93.55 | 94.53 | 98.83 | 98.83 | 99.8 |
| 10 | 98.83 | 93.75 | 94.14 | 99.02 | 85.35 | 90.23 | 95.51 | 97.85 | 96.88 |
| Mean | 97.7 | 97.54 | 95.25 | 99.14 | 93.42 | 93.95 | 97.23 | 97.21 | 98.03 |
| Uncertainty | 4.06 | 2.53 | 3.49 | 1.09 | 5.37 | 4.03 | 2.8 | 4.23 | 1.87 |

**Chart, box and whisker chart

Description automatically generated**

The inception models also perform very well, with the best model achieving 99.14% mean test accuracy.

Chart, waterfall chart

Description automatically generated

**SmecticIF, sequential**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Validation accuracies** | | | | | | |
| Run | Architecture (number of layers, starting channels) | | | | | |
|  | 3, 8 | 3, 16 | 3, 32 | 4, 8 | 4, 16 | 4, 32 |
| 1 | 82.67 | 68.75 | 72.73 | 79.55 | 80.4 | 73.86 |
| 2 | 82.67 | 76.99 | 87.22 | 76.42 | 75.28 | 82.95 |
| 3 | 78.69 | 72.44 | 66.48 | 96.59 | 71.31 | 87.78 |
| 4 | 68.75 | 75.85 | 76.14 | 81.25 | 74.72 | 77.56 |
| 5 | 78.69 | 60.23 | 75.28 | 82.39 | 81.53 | 81.82 |
| 6 | 76.42 | 84.66 | 77.84 | 74.43 | 89.77 | 78.98 |
| 7 | 83.52 | 83.24 | 76.7 | 92.33 | 78.41 | 79.26 |
| 8 | 76.14 | 82.39 | 80.97 | 76.99 | 82.95 | 86.08 |
| 9 | 78.69 | 70.45 | 72.16 | 81.53 | 63.07 | 90.34 |
| 10 | 56.82 | 81.53 | 87.78 | 87.5 | 77.84 | 65.06 |
| Mean | 76.31 | 75.65 | 77.33 | 82.9 | 77.53 | 80.37 |
| Uncertainty | 7.66 | 7.37 | 6.26 | 6.8 | 6.82 | 6.95 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Test accuracies** | | | | | | |
| Run | Architecture (number of layers, starting channels) | | | | | |
|  | 3, 8 | 3, 16 | 3, 32 | 4, 8 | 4, 16 | 4, 32 |
| 1 | 51.63 | 46.74 | 77.99 | 68.75 | 69.57 | 71.74 |
| 2 | 69.84 | 69.29 | 54.62 | 63.59 | 77.72 | 68.75 |
| 3 | 67.93 | 86.96 | 64.95 | 49.46 | 60.6 | 65.76 |
| 4 | 57.88 | 73.1 | 68.48 | 59.24 | 46.47 | 64.13 |
| 5 | 75 | 65.76 | 73.91 | 37.5 | 52.99 | 54.35 |
| 6 | 61.41 | 57.07 | 63.32 | 57.34 | 53.26 | 60.33 |
| 7 | 72.28 | 64.4 | 70.11 | 56.25 | 66.58 | 38.86 |
| 8 | 53.53 | 70.11 | 67.12 | 74.46 | 55.71 | 53.8 |
| 9 | 64.67 | 79.89 | 73.64 | 61.96 | 61.68 | 65.22 |
| 10 | 70.65 | 65.76 | 56.79 | 39.95 | 70.11 | 50.54 |
| Mean | 64.48 | 67.91 | 67.09 | 56.85 | 61.47 | 59.35 |
| Uncertainty | 7.67 | 10.62 | 7.07 | 11.15 | 9.1 | 9.46 |

**Chart, box and whisker chart

Description automatically generated**

We see worse results for the smecticIF set. The validation accuracies fall in the range of 70-80%, with even lower test accuracies, with the best at 67.91%.

Chart, waterfall chart

Description automatically generated

**SmecticIF, inception**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Validation accuracies** | | | | | | | | | |
| Run | Architecture (number of blocks, starting channels) | | | | | | | | |
|  | 1, 2 | 1, 4 | 1, 8 | 2, 2 | 2, 4 | 2, 8 | 3, 2 | 3, 4 | 3, 8 |
| 1 | 68.47 | 82.67 | 80.97 | 82.67 | 83.52 | 79.26 | 86.08 | 86.93 | 92.9 |
| 2 | 74.43 | 71.88 | 85.8 | 80.97 | 89.49 | 83.24 | 76.14 | 71.59 | 90.34 |
| 3 | 53.98 | 76.14 | 74.72 | 54.55 | 54.55 | 76.42 | 86.36 | 75.85 | 83.81 |
| 4 | 76.99 | 92.05 | 84.94 | 78.98 | 76.99 | 96.59 | 92.61 | 77.84 | 90.06 |
| 5 | 90.34 | 86.65 | 94.6 | 72.73 | 63.64 | 77.84 | 88.92 | 73.3 | 87.5 |
| 6 | 46.02 | 73.58 | 84.09 | 93.18 | 89.49 | 82.39 | 79.26 | 90.91 | 80.97 |
| 7 | 45.17 | 93.75 | 98.01 | 70.17 | 81.53 | 94.03 | 69.03 | 92.05 | 81.82 |
| 8 | 82.39 | 91.19 | 92.9 | 75 | 83.24 | 84.38 | 83.81 | 73.01 | 95.74 |
| 9 | 66.76 | 86.65 | 95.17 | 84.38 | 57.39 | 89.77 | 90.62 | 54.83 | 98.58 |
| 10 | 84.94 | 85.23 | 80.68 | 89.77 | 86.36 | 92.9 | 79.26 | 88.35 | 88.92 |
| Mean | 68.95 | 83.98 | 87.19 | 78.24 | 76.62 | 85.68 | 83.21 | 78.47 | 89.06 |
| Uncertainty | 15.2 | 7.39 | 7.22 | 10.44 | 12.52 | 6.82 | 6.9 | 10.85 | 5.48 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Test accuracies** | | | | | | | | | |
| Run | Architecture (number of blocks, starting channels) | | | | | | | | |
|  | 1, 2 | 1, 4 | 1, 8 | 2, 2 | 2, 4 | 2, 8 | 3, 2 | 3, 4 | 3, 8 |
| 1 | 63.32 | 64.67 | 57.61 | 63.32 | 52.45 | 58.15 | 53.8 | 43.48 | 46.2 |
| 2 | 76.9 | 57.07 | 58.42 | 57.61 | 65.22 | 59.51 | 38.86 | 60.05 | 43.21 |
| 3 | 57.61 | 63.86 | 66.85 | 56.25 | 56.52 | 61.41 | 58.42 | 56.79 | 57.61 |
| 4 | 48.91 | 66.58 | 52.45 | 50.82 | 48.37 | 60.87 | 52.99 | 63.32 | 42.66 |
| 5 | 44.57 | 64.13 | 50.54 | 60.87 | 58.7 | 64.67 | 55.98 | 56.79 | 44.29 |
| 6 | 44.02 | 46.47 | 64.95 | 41.85 | 41.85 | 65.49 | 71.74 | 43.48 | 61.96 |
| 7 | 72.01 | 49.18 | 52.17 | 67.66 | 57.88 | 47.28 | 45.92 | 54.62 | 60.05 |
| 8 | 50.27 | 51.36 | 49.18 | 42.12 | 56.52 | 38.86 | 58.42 | 33.7 | 46.47 |
| 9 | 72.55 | 59.24 | 48.64 | 54.08 | 66.03 | 60.33 | 54.35 | 56.52 | 38.04 |
| 10 | 67.66 | 58.42 | 49.73 | 48.91 | 59.78 | 53.26 | 50.54 | 47.83 | 54.89 |
| Mean | 59.78 | 58.1 | 55.05 | 54.35 | 56.33 | 56.98 | 54.1 | 51.66 | 49.54 |
| Uncertainty | 11.72 | 6.69 | 6.28 | 8.14 | 6.94 | 7.88 | 8.16 | 8.73 | 7.91 |

**Chart, box and whisker chart

Description automatically generated**

The inception models perform even worse on the test set than the sequential models.

Chart, waterfall chart

Description automatically generated

Further investigation or expansion of the smecticIF dataset is needed.