Pepper Seed Classification with Neural Networks

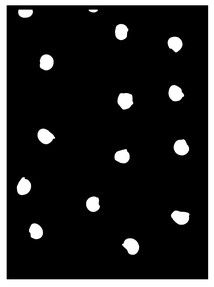
Author: Jon Eman
Department: Computer Science

Image Processing

The dataset was carefully acquired and processed using a Canon LiDE 300 Flatbed Scanner and OpenCV

Seeds were scanned in bulk at 2400dpi, blurred with a 15px X 15px kernel to remove noise, and a 100px border was added

The image was then thresholded on the red channel at an intensity level of 100, morphologically closed with a 25px X 25px kernel to fill any gaps, and then blurred with a 9px X 9px kernel to accommodate edge detection



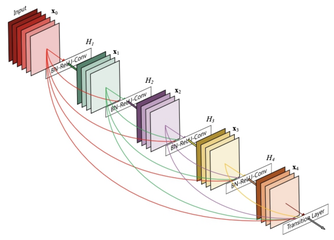
Canny edge detection was applied, and contours were drawn for each seed on its own threshold image. Each threshold was applied to the original image, to produce an separate image for each individual seed. If the resulting seed image had a width or height of less than 100px, it was removed from the dataset. Otherwise, the image was padded with 0 intensity pixels to make each seed image 500px X 500px

Model Training

The best pre-built image classification models were trained against the data set of 1922 total images of eight cultivars of pepper seeds, to choose the best base for the final model

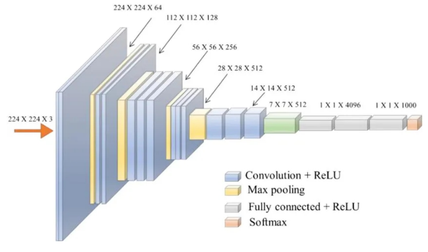
Model	Size (MB)	ImageNet Top-1 Accuracy	Parameters	Depth	Pepper Seed Accuracy
Xception	88	79.0%	22.9M	81	72.8%
VGG16	528	71.3%	138.4M	16	76.3%
VGG19	549	71.3%	143.7M	19	75.2%
ResNet50	98	74.9%	25.6M	107	47.2%
ResNet50V2	98	76.0%	25.6M	103	65.9%
ResNet101	171	76.4%	44.7M	209	49.1%
ResNet101V2	171	77.2%	44.7M	205	69.4%
ResNet152	232	76.6%	60.4M	311	50.3%
ResNet152V2	232	78.0%	60.4M	307	71.1%
InceptionV3	92	77.9%	23.9M	189	73.0%
InceptionResNetV2	215	80.3%	55.9M	449	68.4%
MobileNet	16	70.4%	4.3M	55	74.8%
MobileNetV2	14	71.3%	3.5M	105	73.0%
DenseNet121	33	75.0%	8.1M	242	76.1%
DenseNet169	57	76.2%	14.3M	338	70.9%
DenseNet201	80	77.3%	20.2M	402	76.1%
NASNetMobile	23	74.4%	5.3M	389	66.3%
NASNetLarge	343	82.5%	88.9M	533	61.9%
EfficientNetB0	29	77.1%	5.3M	132	14.3%

DenseNet201 Architecture



Source: https://doi.org/10.48550/arXiv.1608.06993

VGG16 Architecture



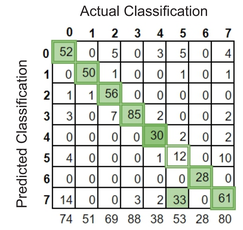
Source: https://doi.org/10.1016/j.rcim.2019.01.006

The biggest difference between the top two models is that DenseNet propagates every input from every layer as an input to every other layer, while VGG16 consolidates and down samples through the layers, until the number of outputs is equal to the number of classes.

Results Evaluation

The top two pre-built models were fine-tuned by unfreezing layers of the models and retraining at very low learning rates.

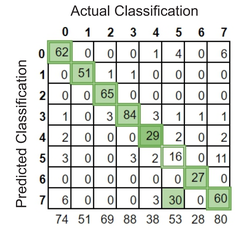
DenseNet201



Classification Accuracy

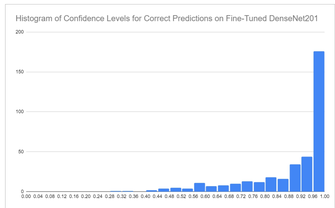
0	Hungarian	70.3%
1	Habanero	98.0%
2	Cayenne	81.2%
3	Serrano	96.6%
4	Poblano	78.9%
5	Jalapeno	22.6%
6	Cherry	100.0%
7	Banana	76.3%
Overall		**77.8%**

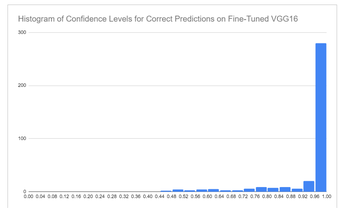
VGG16

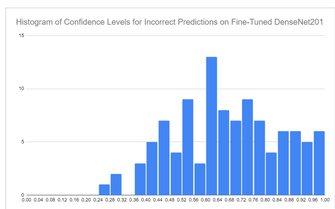


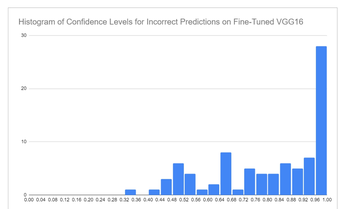
Classification Accuracy

0	Hungarian	83.8%
1	Habanero	100.0%
2	Cayenne	94.2%
3	Serrano	95.5%
4	Poblano	76.3%
5	Jalapeno	30.2%
6	Cherry	96.4%
7	Banana	75.0%
Overall		**81.9%**









UMSL | University of Missouri–St. Louis