

Playing Card Classifier

Classify colour images of playing cards into 53 different classes.

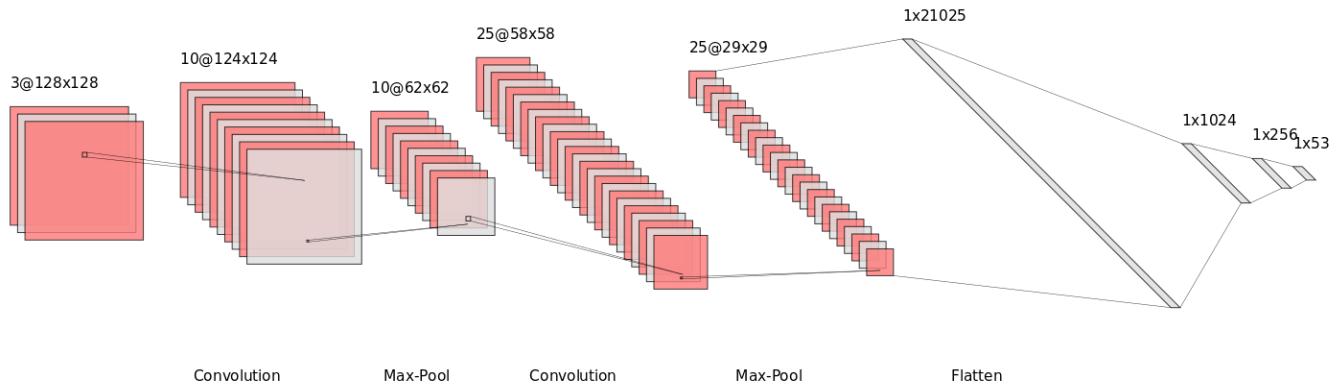
Learning Objectives

- Image classification
- Convolutional Neural Networks (CNN)
- ANN architecture tuning
- Image manipulation and distortion
- Training with validation
- Batch training

Visualising the Data



Convolutional Neural Network



A CNN was constructed, taking in a 3-channel 128x128 image and outputting a soft-max normalised vector of length 53, giving the probability for each class. The network consisted of two 2D convolution layers (kernel size = 5), each followed by max-pool layers (kernel size = 2), and four linear layers (sizes = 21025, 1024, 256, 53).

The model was trained in batches of 32 images using the Adam optimiser until the cross-entropy loss on a validation set converged (~8 epochs). The validation set was used to prevent overfitting on the training dataset. The model achieved a score of 0.9791 on the training dataset and 0.7925 on the test dataset. The model was most confused by number cards displaying pictures in their centre.

Architecture Tuning

Batch Size = 16, 32, 64

MaxPoolSize = 2, 4, 6

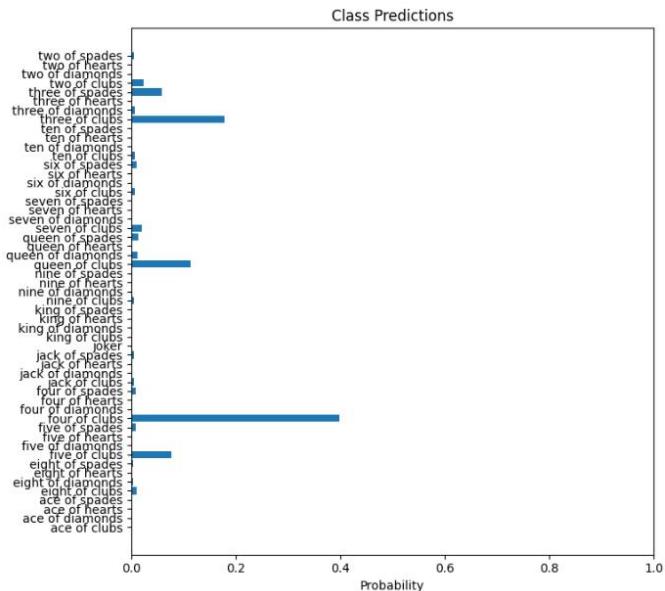
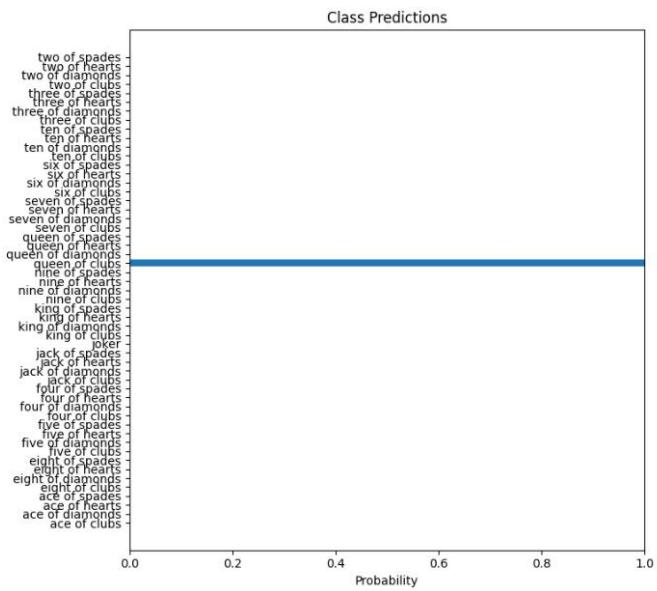
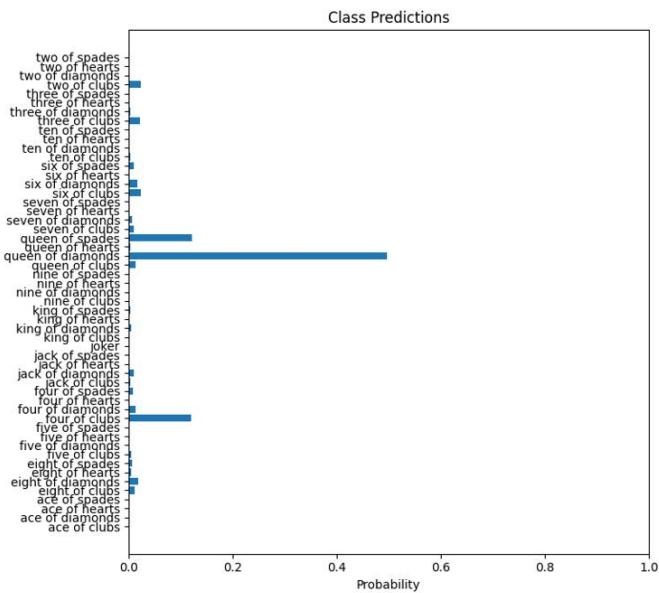
KernelSize = 3, 5, 8

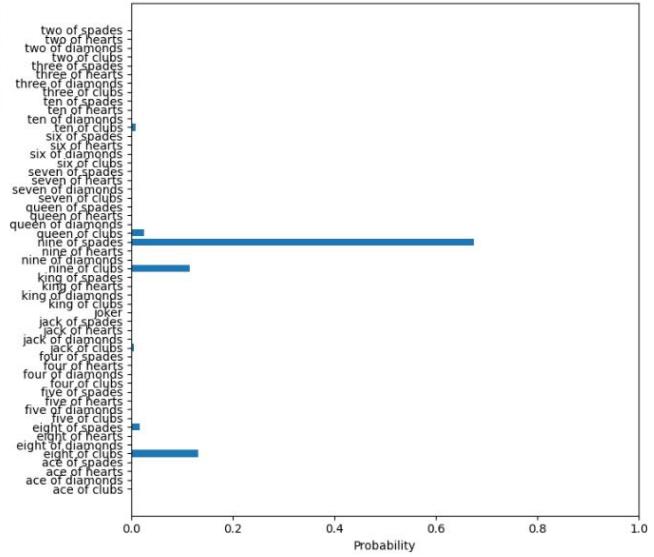
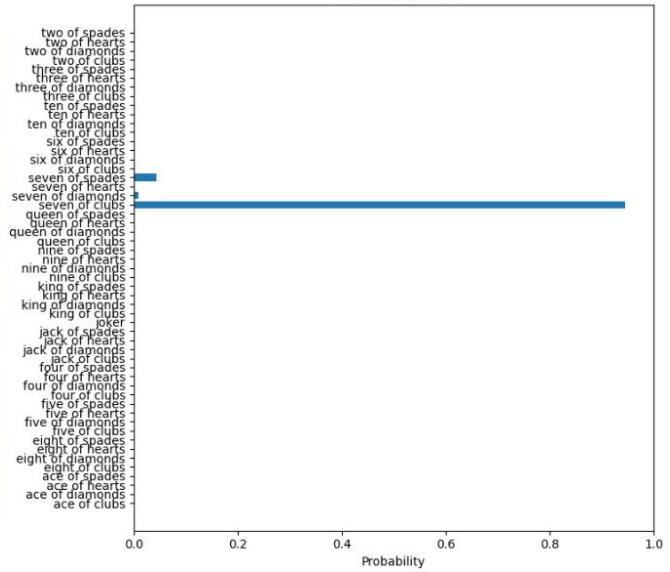
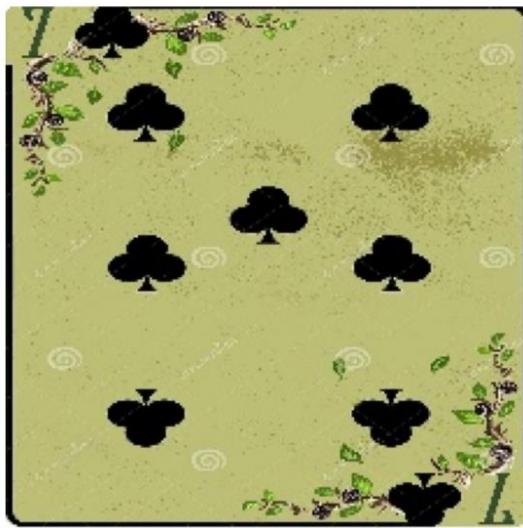
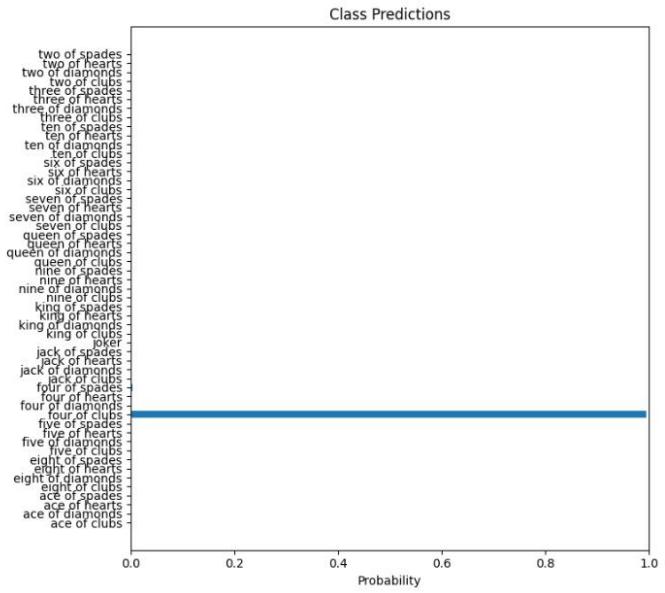
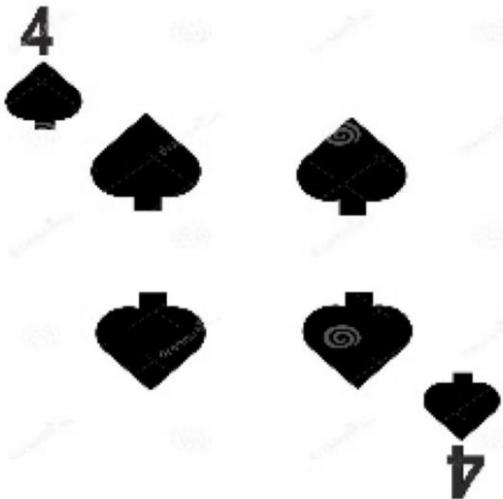
KernelLayers = 5, 10

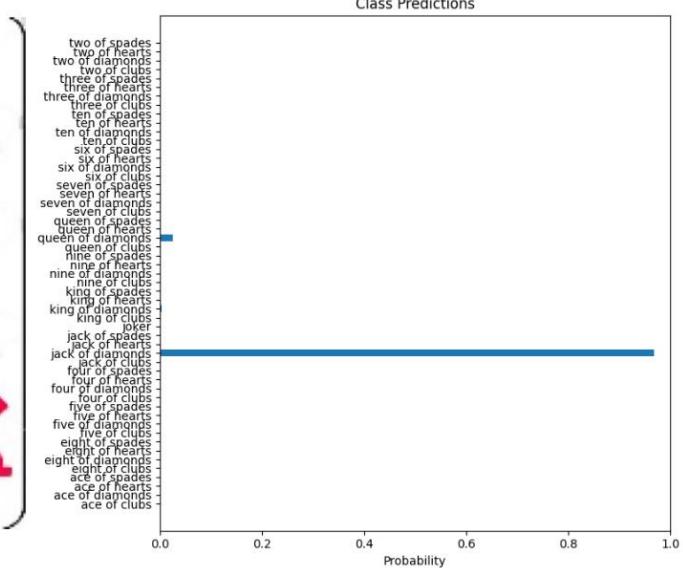
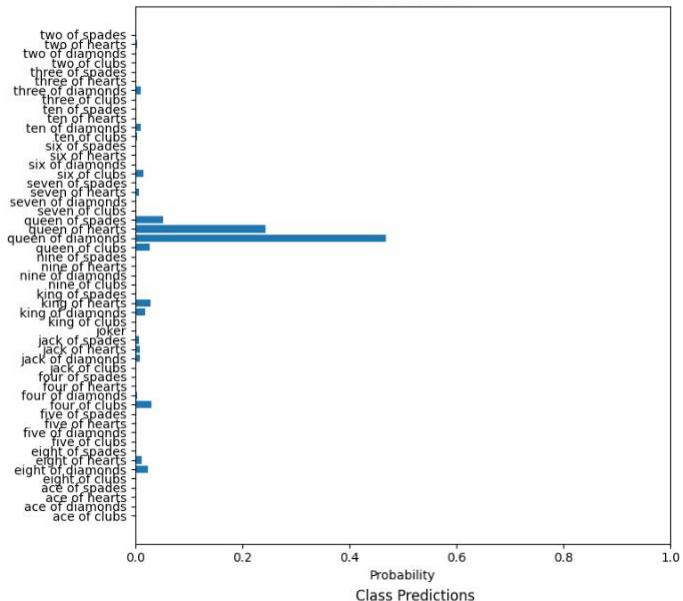
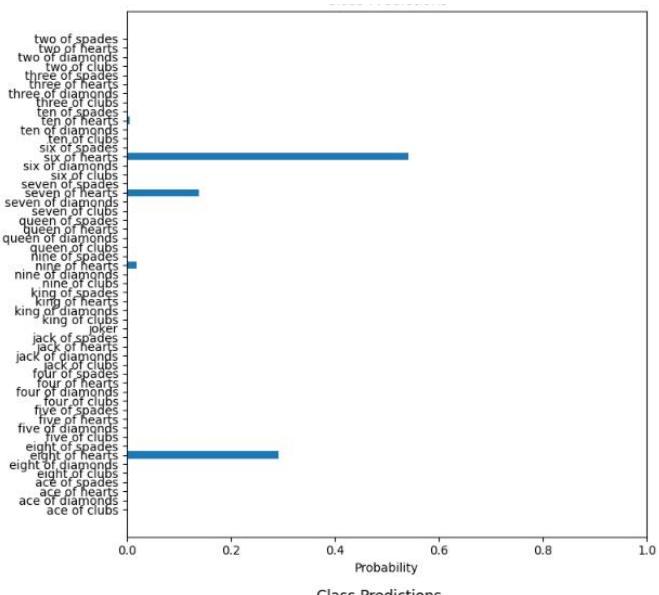
LinearSize = 1024, 8192

MaxPoolSize	KernelSize	KernelLayers	LinearSize	Val Score
4	5	10	1024	0.79245
2	8	5	8192	0.78491
6	8	10	1024	0.69057
4	3	5	8192	0.77736
2	3	5	1024	0.76603
6	5	10	8192	0.73208
4	5	5	1024	0.78868
2	3	10	8192	0.76604
6	3	5	1024	0.70566
6	8	5	8192	0.69811
4	3	10	1024	0.77358
2	5	10	8192	0.80000
4	5	10	8192	0.81509

Visualising Performance







Conclusion

The convolutional neural network was very capable at differentiating playing cards, even when the image was distorted. The convolution and max pooling layers identified patterns associated with different ranks and suits. Medium-sized kernels produced the best results – too small kernels would not capture enough spatial information, whilst too large kernels would miss finer details. Having more kernels and a larger fully connected layer allowed the model to handle more information, also producing more accurate results.

The model appears to be most confused by number cards with pictures in the centre (mistaking them for picture cards). It also sometimes mistakes the spade symbol for clubs, and vice versa, suggesting that the model recognises the colour and shape of these symbols.