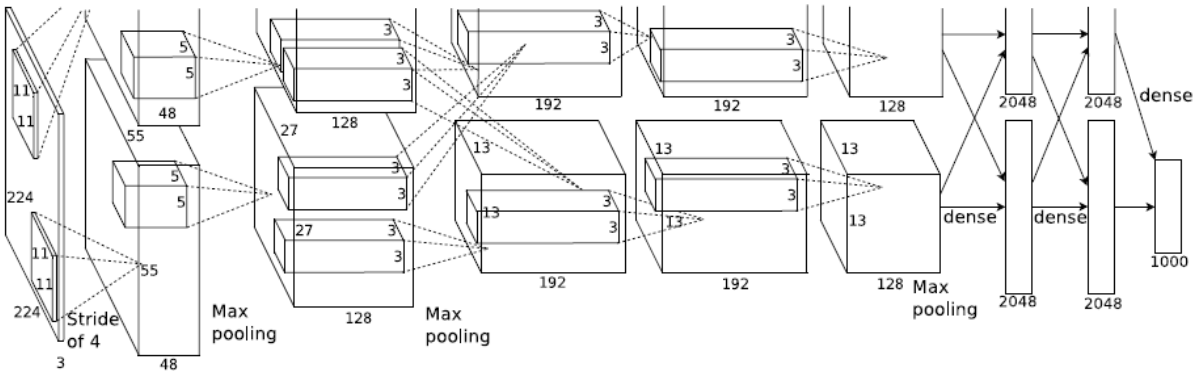
A. Using pretrained CNNs to classify a single image

In this part, a pretrained convolutional neural network is used to classify an image. The CCN that was used is AlexNet, which had been trained on a subset of the ImageNet database, which was used in ImageNet Large-Scale Visual Recognition Challenge (ILSVRC). AlexNet is trained on more than a million images to classify 1000 classes. AlexNet architecture is shown in the figure below.



The CNN is stored using command net = AlexNet and then the image is cropped into the required input size of the AlexNet, then the image is classified using the command classify (net, I).
The image was classified correctly.

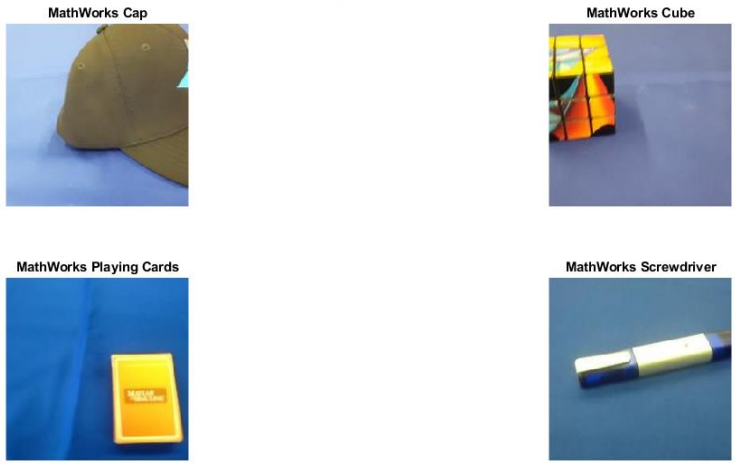
Pretrained VGG-16 network model for image classification
The same experiment as above was done using the VGG16 CNN model. The model classified the input image accurately.

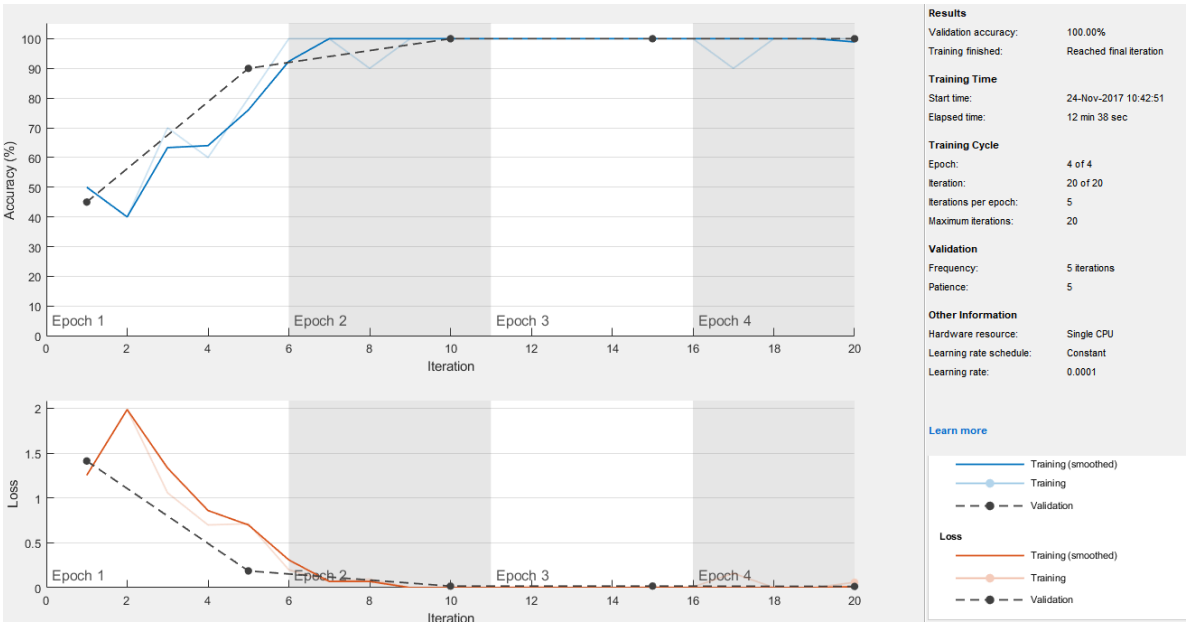
B. Transfer learning

1.

The pretrained AlexNet CNN is used to perform classification on a new set of images instead of the original ones. AlexNet is a CNN and had been trained on over one million images in order to classify 1000 types of objects. Transfer learning concept is used in this part to retrain the AlexNet on order to classify a new set of images into 5 different categories. In order to do transfer learning, the extracted features in the first layers of the network will be used and the final layers of the network will be retrained to adapt to the new classification task. Therefore, the last three layers of the AlexNet will be replaced with
a fully connected layer, a softmax layer, and a classification output layer. And the fully connected layer will be set to have 5 outputs instead of 1000 outputs. The learning rate of the initial layers will be set to a small value and the learning rate of the new layers will be set to a high value in order to keep the original weights of the transferred layers while adjusting the last layers for the new task. Transfer learning is a faster process compared to original training, because there is no need to train on many epochs. After specifying the layers and options for training, the training can be done with the command trainNetwork in the matlab. The new transferred network can be used to classify the validation set images by using

command classify, and the accuracy was 100%. The following figure shows four sample validation images with their predicted labels.





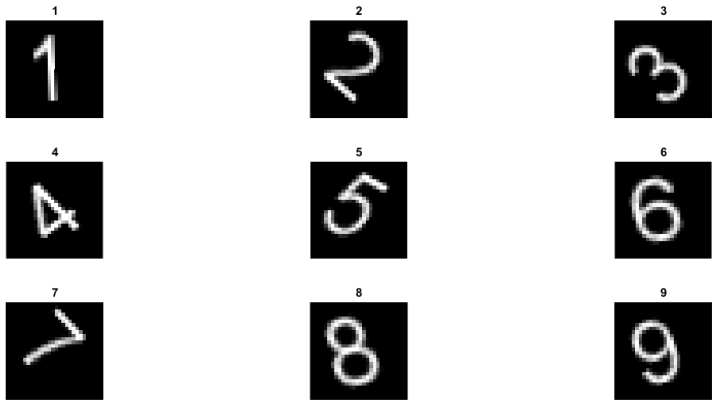
2.

In this part, we load a pretrained network from the file LettersClassificationNet.mat that was originally trained on a large collection of 28-by-28 grayscale letter images. This network classifies images into 3 letter classes of ‘A’, ‘B’ and ‘C’. The net has the following 7 layers:

1 'imageinput' Image Input 28x28x1 images with 'zerocenter' normalization
2 'conv' Convolution 20 5x5x1 convolutions with stride [1 1] and padding [0 0 0 0]

3	'relu'	ReLU	ReLU
4	'maxpool'	Max Pooling	2x2 max pooling with stride [2 2] and padding [0 0 0 0]
5	'fc'	Fully Connected	3 fully connected layer
6	'softmax'	Softmax	softmax
7	'classoutput'	Classi. Output	crossentropyex with 'A', 'B', and 1 other classes

We are going to use this net with transfer learning to classify digits instead of letters. Hence, we replace the last 3 layers of the net with new layers and retrain the network with digit images. The dataset that we use has 10000 synthetic digit images, which half of them is used for retraining the net and half is used for testing the resulted net. We set a high value for the learning rate of the fully connected layer and a very small learning rate for the transferred layers in order to train the network while keeping the original features.
After fine-tuning the network, the test image set was classified using it and the classification accuracy was 0.9160.The following figure shows a sample of test images along with their predicted class.



C. Using a pretrained CNN as a feature extraction block

In this part, the feature of the images are extracted using a pretrained CNN and they are used for classification using support vector machine (SVM). The dataset contains 75 images, which from it 70% is used for training and 30% for testing. AlexNet is used as the CNN for feature extraction. Training features were extracted by using command activations (net,trainingImages,layer) with the layer 'fc7'. The extracted features from training images were used to fit a multiclass support vector machine (SVM) using the command fitcecoc. The classification accuracy on the test set was 95%.The following figure shows a sample of test images with their predicted class.

MathWorks Cap

MathWorks Cube

MathWorks Playing Cards

MathWorks Screwdriver