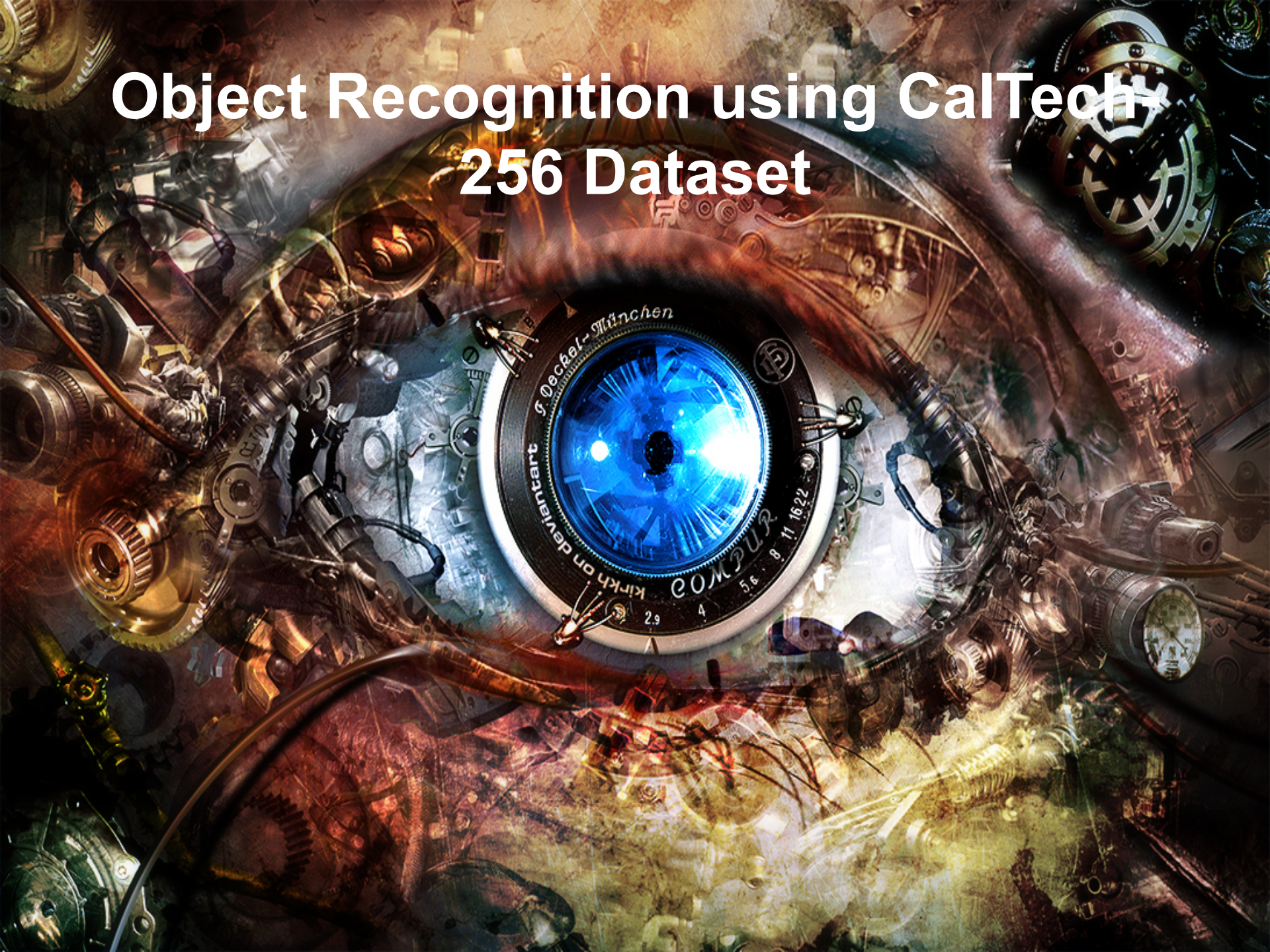


Object Recognition using CalTech-256 Dataset

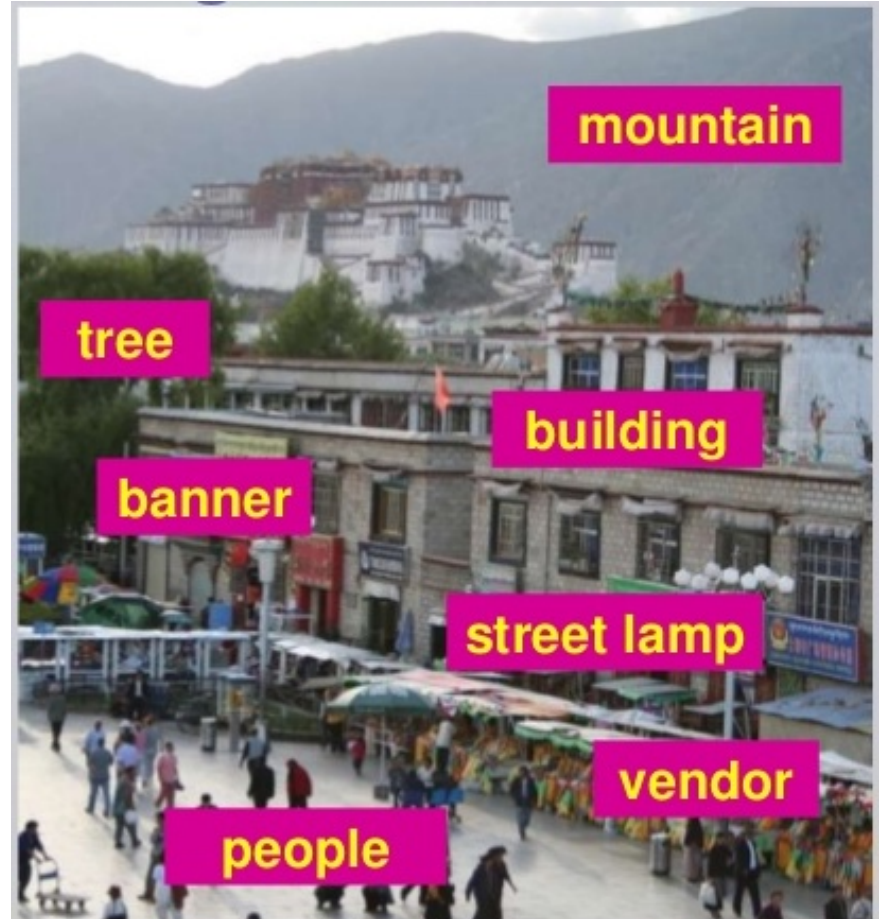


Introduction



AnubhabMajumdar, ShubhamMunot,
ToshalPhene

Introduction



Problem statement



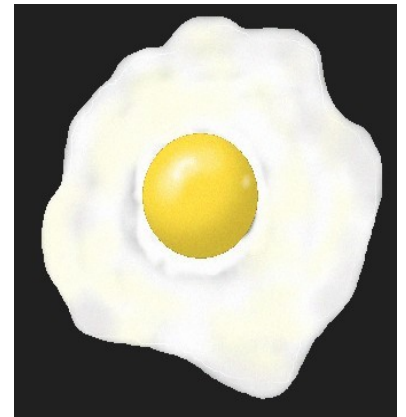
Bag



Binocular



Eiffel Tower



Fried egg

Challenges



Large variety

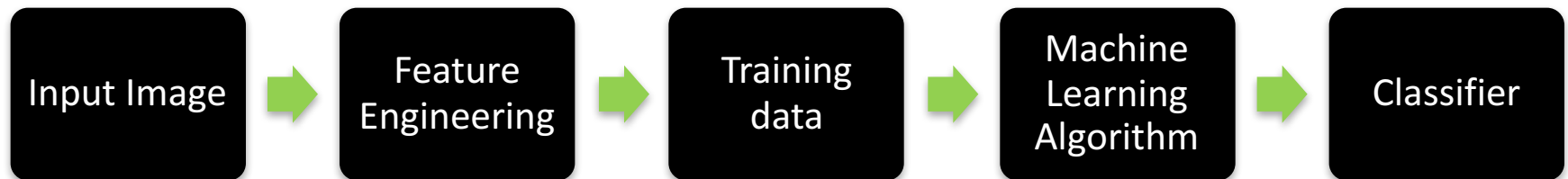


Clutter,
illumination



Scale, orientation,
occlusion

Traditional approach

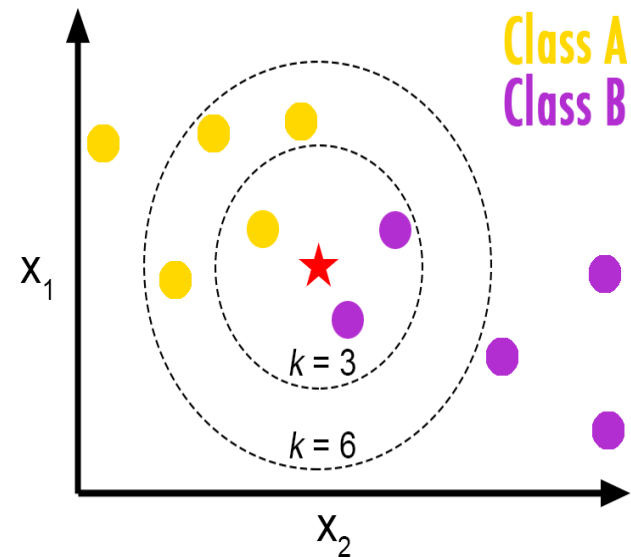


Feature engineering

- Gabor filters
 - Provides texture of an image
 - Similar to how human distinguish texture

Method 1: KNN

- Important parameters
 - Distance measure
 - Suitable value of K



Sample representation

KNN: Results

k	Distance measure	Test accuracy
1	euclidean	55.71
5	euclidean	56.9755
25	euclidean	51.9138
50	euclidean	53.1792
1	cosine	54.4446
5	cosine	55.71
25	cosine	50.6483
50	cosine	48.1175
1	cityblock	55.71
5	cityblock	60.7718
25	cityblock	54.4446
50	cityblock	51.9138
1	correlation	54.4446
5	correlation	55.71
25	correlation	50.6483
50	correlation	48.1175

Method 2: SVM

- Important parameters
 - Kernel (linear, polynomial, Radial Basis Function, sigmoid)
 - C (error penalty) and gamma (relax constraint)

SVM: Results

c	gamma	Kernel	Train accuracy	Test accuracy
128	3.05175E-05	linear		42.50%
128	3.05175E-05	poly	42.80%	25%
128	3.05175E-05	RBF	66.59%	56.25%
128	3.05175E-05	sigmoid	62.22%	57.50%

Method 3: ANN

- Method 1 : Gabor filter
- Method 2 : Linearization
- Hidden layers use sigmoid activation function
- Neurons in output layer use softmax activation function

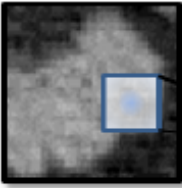
ANN: Results

Method	Input Size	Hidden layer	# of neurons in hidden layer	Train	Test
Gabor filter	4096	1	2048	81.10%	50.60%
Gabor filter	4096	1	4096	63.60%	54.30%
Gabor filter	4096	1	8192	81.10%	50.60%
Linear	784	1	1024	99%	70.40%
Linear	784	1	2048	92%	71%

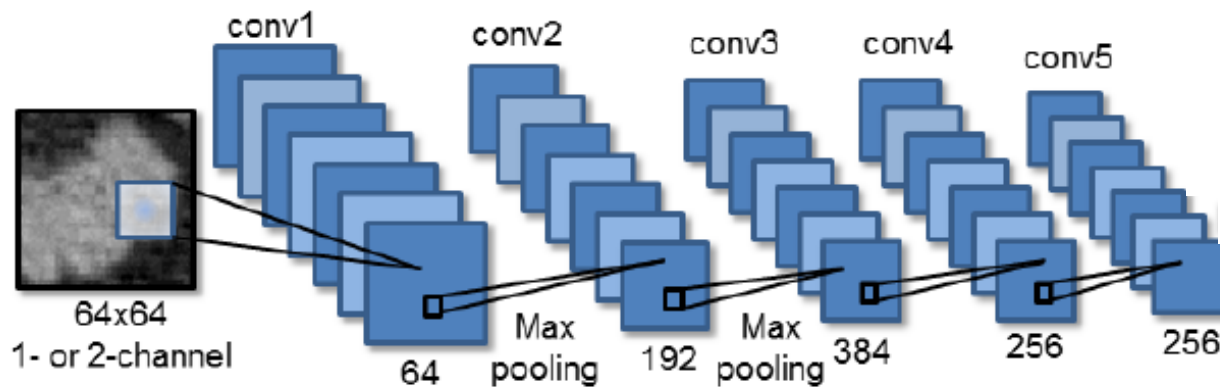
A State of the art technique : Deep Convolutional Neural Network

- Popularized by LeCun et. al. in their pioneering work LeNet-5, a 7-level network by that classifies digits
- A convolutional neural network (CNN, or ConvNet) is a type of feed-forward artificial neural network (ANN)
- The connectivity pattern between its neurons is inspired by the organization of the animal visual cortex

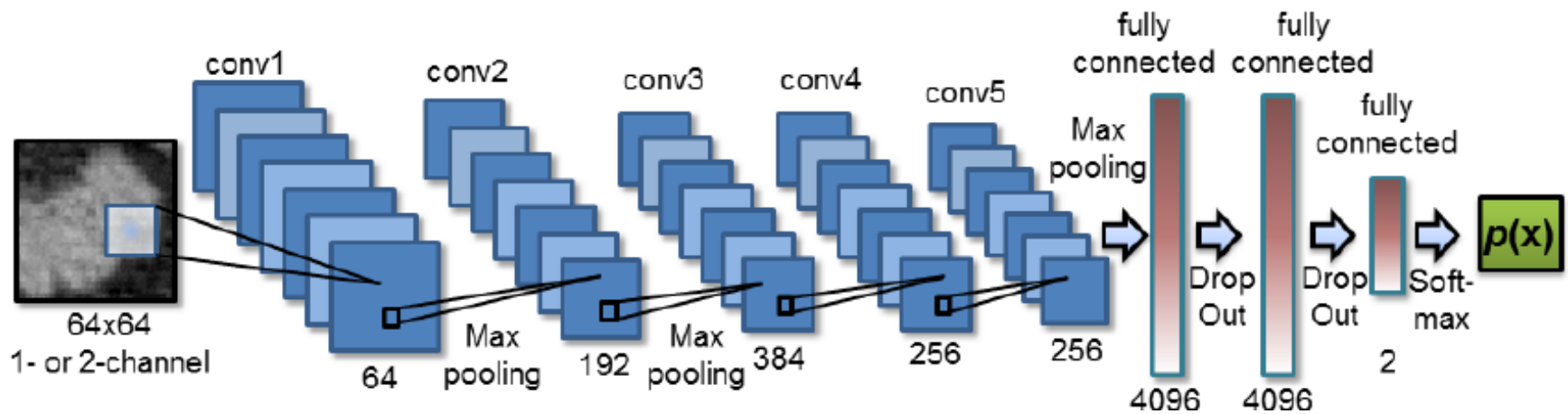
Typical ConvNet architecture



Typical ConvNet architecture



Typical ConvNet architecture

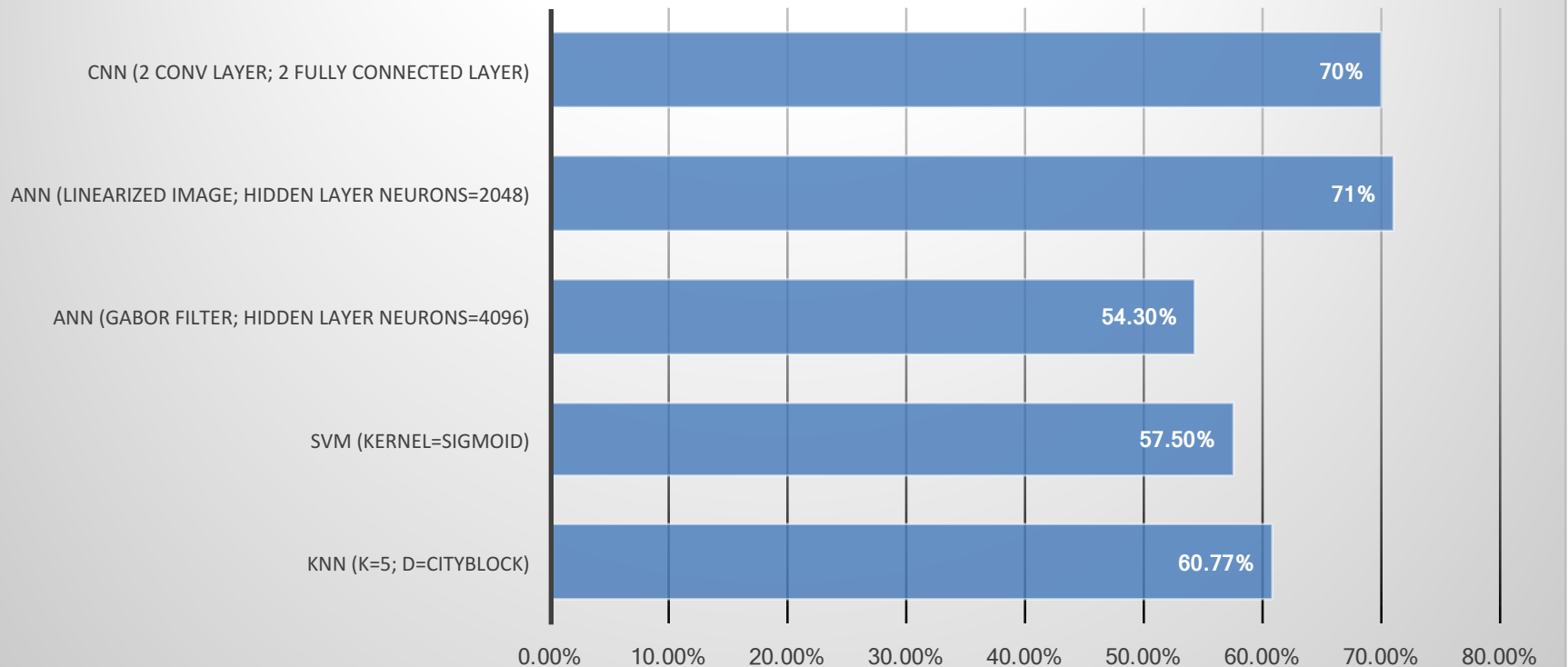


Results

Index	Input Image Size	# of conv layer	# of features in conv layer 1	# of features in conv layer 2	# of features in conv layer 3	# of fully connected layer	Fully connected layer 1	Fully connected layer 2	Train Accuracy	Test Accuracy
CNN 1	28x28	2	32	64	NA	1	64	NA	45.43%	26.25%
CNN 2	28x28	2	32	64	NA	1	1024	NA	79.78%	27.50%
CNN 3	64x64	3	32	64	128	1	1024	NA	91%	42.50%
CNN 4	28x28	2	64	128	NA	2	1024	1024	96.82%	69.69%
CNN 5	28x28	2	128	256	NA	2	1024	1024	96.54%	67.81%

Comparison

Test accuracy



Conclusion

- Handcrafted features for large variety of images is unreliable and fragile

Conclusion

- Handcrafted features for large variety of images is unreliable and fragile
- Featureless approach to computer vision problem is more robust

Conclusion

- Handcrafted features for large variety of images is unreliable and fragile
- Featureless approach to computer vision problem is more robust
- Deep convolutional neural nets trained with large dataset yield better accuracy than learning algorithm with traditional hand crafted feature (Ref: ImageNet)

Conclusion

- Handcrafted features for large variety of images is unreliable and fragile
- Featureless approach to computer vision problem is more robust
- Deep convolutional neural nets trained with large dataset yield better accuracy than learning algorithm with traditional hand crafted feature (Ref: ImageNet)
- Modern computer vision research is moving towards featureless approach with ConvNet
 - ImageNet
 - GoogLeNet

References

- Zhang et. al. - SVM-KNN: Discriminative Nearest Neighbor Classification for Visual Category Recognition
- Lecun et. al. - Backpropagation Applied to Handwritten Zip Code Recognition
- Krizhevsky et. al. - ImageNet Classification with Deep Convolutional Neural Networks
- Zeiler et. al. - Visualizing and Understanding Convolutional Networks
- [Deep MNIST for Experts](#) using TensorFlow