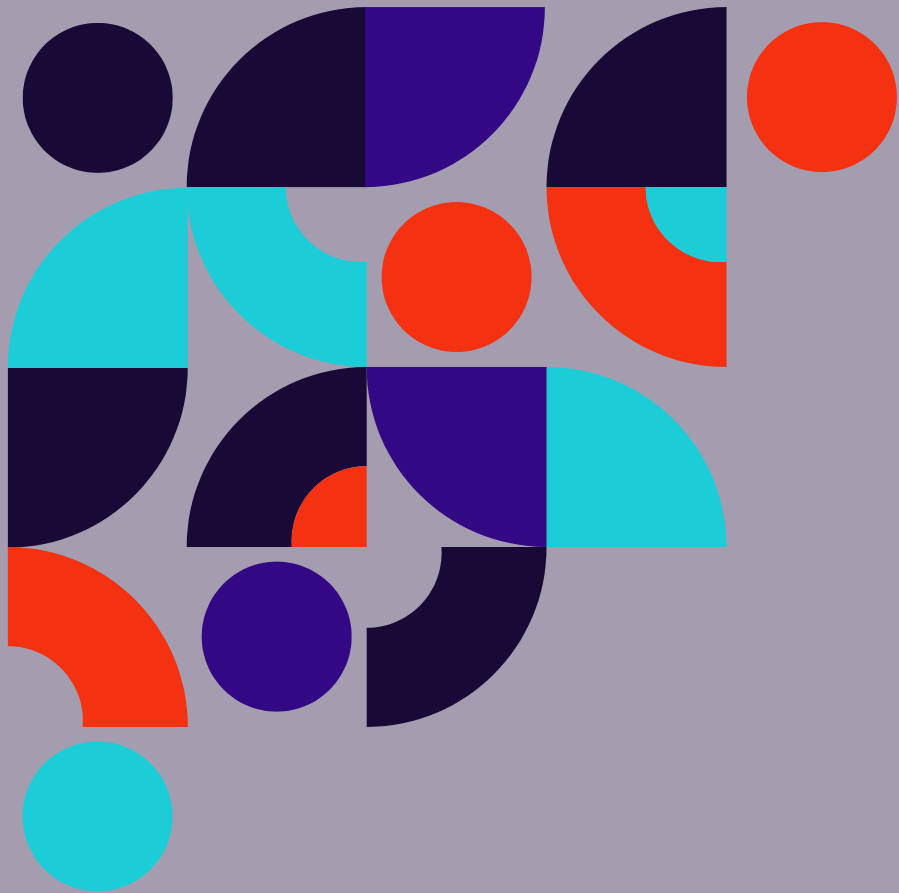


The background of the slide is a high-quality photograph of fresh berries. On the left side, there is a dense cluster of strawberries with bright red, textured skin and green leafy tops. Interspersed among the strawberries are several blueberries, which are small, round, and have a deep blue color with a slight white bloom. There are also raspberries, showing their characteristic bumpy, red structure, and a few blackberries, which are dark purple-black and composed of many small drupelets. The berries are resting on a light-colored wooden surface with a prominent, cracked, and weathered texture, possibly a wooden board or a piece of bark.

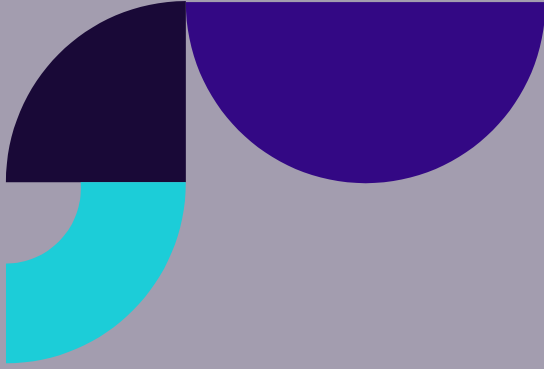
Fruit Image Classification

*Rebecca Brent
Jinghan Fu*



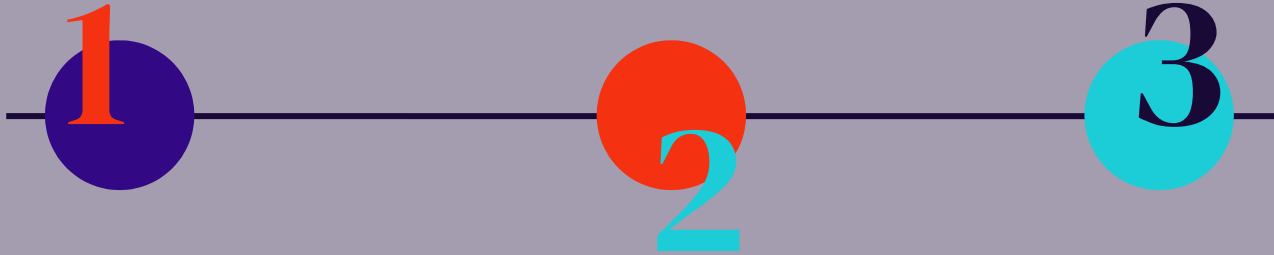
About the Dataset

- Fruits-360 Dataset from Kaggle
- Total number of images: 20118
- Number of classes: 31
- Image size : 100 X 100 pixels



Proposed Questions

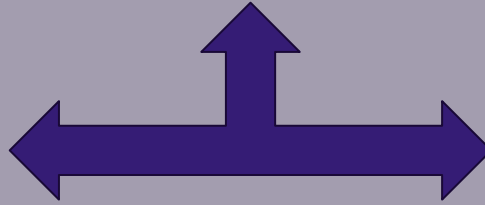
By randomly choosing only a single angle from the train data for each class, how accurate is the model prediction for all test angles?



Can a single piece of produce be reliably classified based on image alone?

How reliably can we classify an image with multiple pieces of fruit?

Data Preprocessing



High-dimensionality

- Each images is 100x100, and each pixel has R, G and B variables: each image has **30000** features

Principal Components Analysis (PCA)

- Outperform other methods in an image recognition task if the number of samples per class is relatively small

Choose the number of components in PCA

- The scree plot shows us that we can keep a lot of variance with very few dimensions.
- Reducing the dataset to 50 components

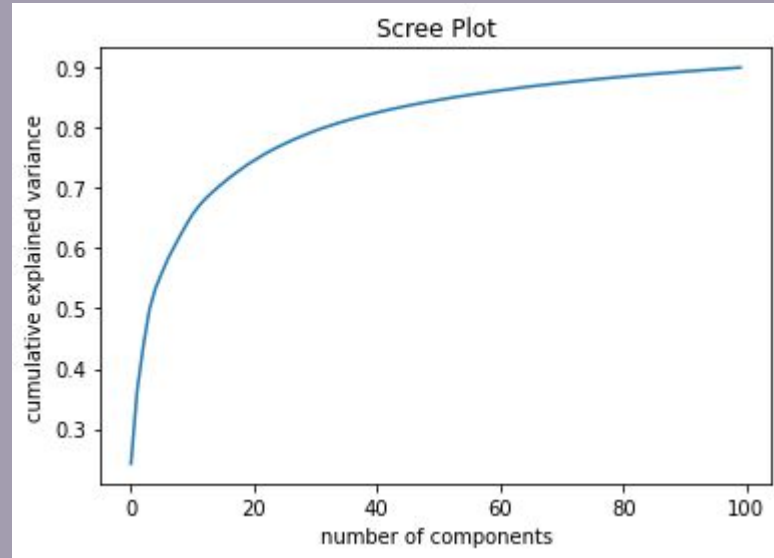
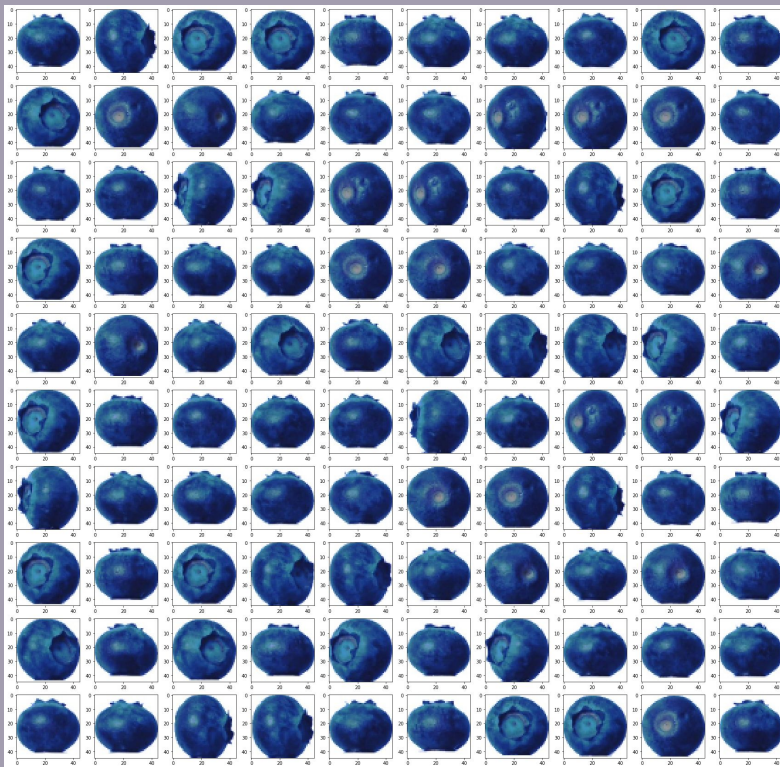


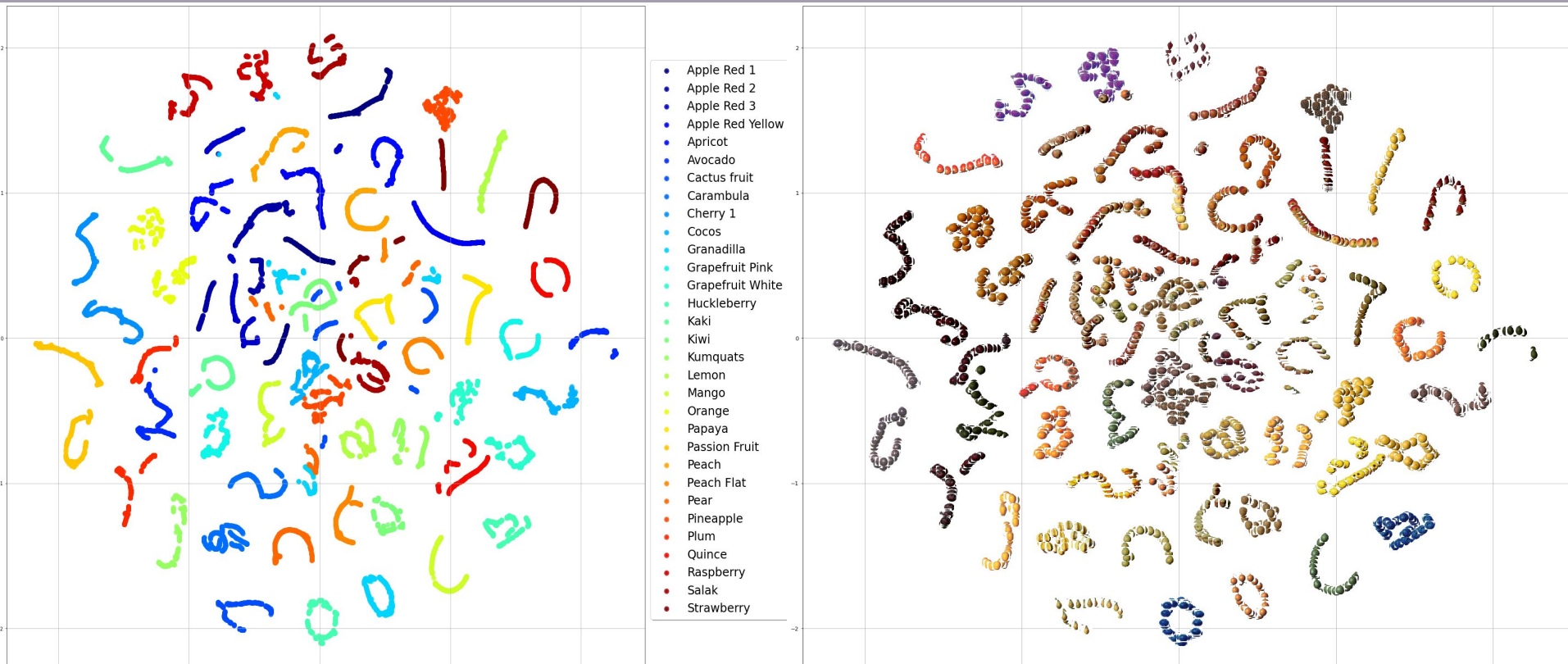
Image Visualization

Sample of 100 original images of huckleberry



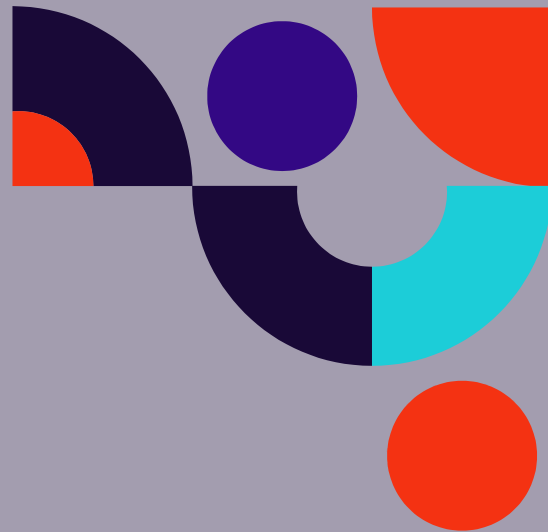
Visualization with PCA(50 components)

Visualizing all 50 PCA components obtained in 2 dimensions.



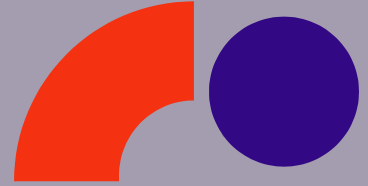
1

*Can a single piece of
produce be reliably
classified based on
image alone?*



Models Results

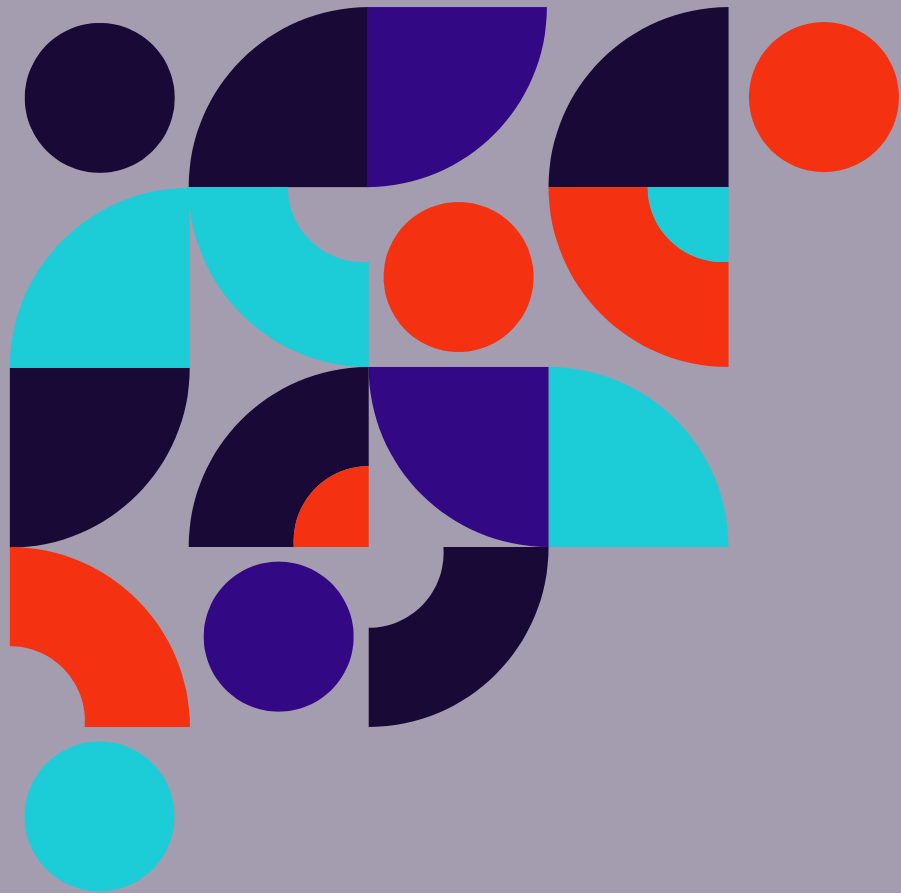
Model	Validation Accuracy	
	Decision Tree	79.47%
	Random Forest	91.04%
	SVM	99.02%
	KNN	96.20%



2

By randomly choosing only a single angle from the train data for each class, how accurate is the model prediction for all test angles?





QUESTION: By randomly choosing only a single angle from the train data for each class, how accurate is the model prediction for all test angles?

Why and How?

- Significantly speed up data collection and computation time
- 1 file path from each image directory chosen
- 30 iterations

Single Angle Results

Random Forest

23%

Decision Tree

17%

**Convolution
Neural Networks**

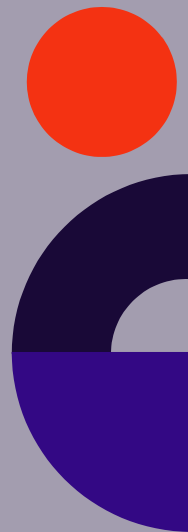
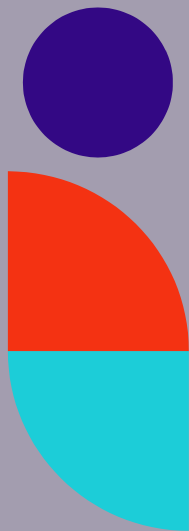
5%

**K Nearest
Neighbors**

39%

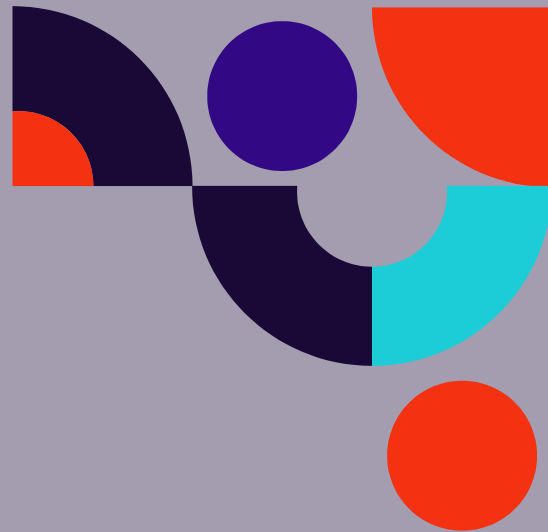
**Support Vector
Machines**

39%



2a

Follow Up: Given that one angle has a very low prediction accuracy how many angles of train data are required to get similar test accuracy to the baseline model which uses all the angles?





QUESTION: How many angles of train data are required to get similar test accuracy to the baseline model which uses all the angles?

Why and How?

- Significantly speed up data collection and computation time
- 2-10, 20, 40, 80 file paths from each image directory chosen
- 30 iterations

Minimum Angles Results

Number of Randomly
Chosen Images per Class

2

10

80

Random Forest

39%

73%

87%

SVM

52%

80%

95%

CNN

8%

31%

91%

KNN

52%

77%

93%

Decision Tree

30%

51%

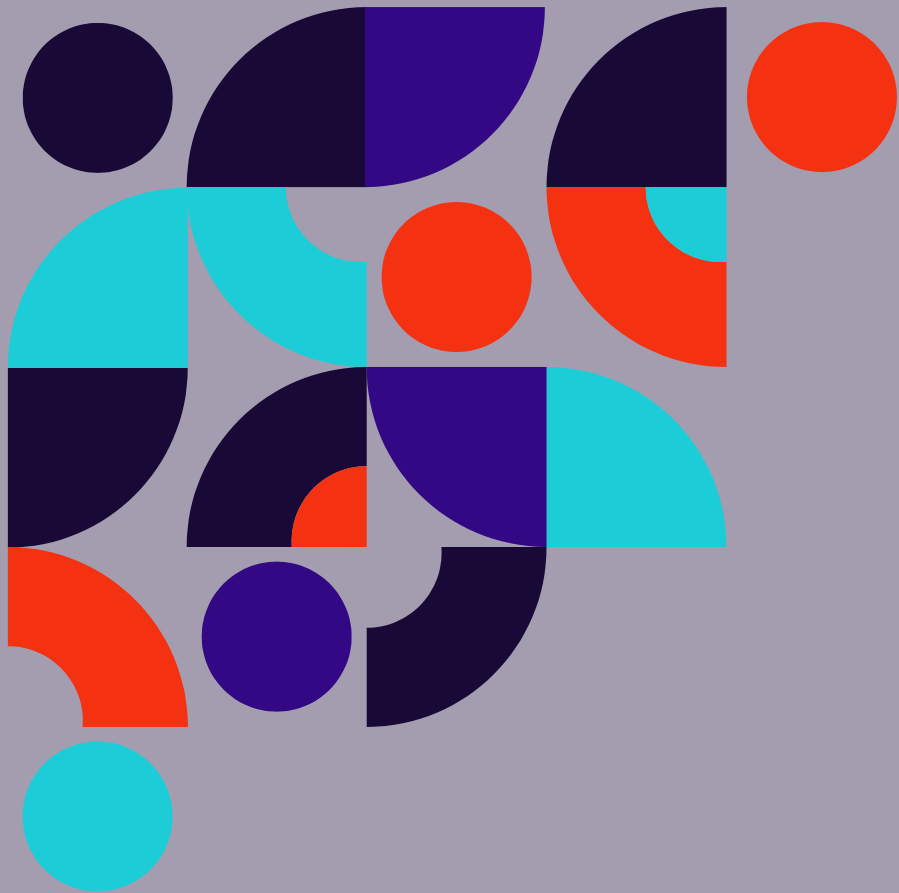
64%

Model

3

*How reliably can we
classify an image with
multiple pieces of fruit?*





QUESTION: How reliably can we classify an image with multiple pieces of fruit?

Why and How?

- Real World Applications
- Convolution Neural Network
- Compressed Apples and Peaches into Single Class

Transformed Image Results



1. Apple, Carambula
2. Cactus, Carambula
3. Carambula

Actual Answer:

Apple, Apricot, Peach,
Pear, Plum



1. Strawberry
2. Strawberry
3. Apple

Actual Answer:

Strawberry



Test Accuracy

1. 4%
2. 6%
3. 4%

Original Image Results



1. No Guesses
2. Cherry, Huckleberry, Strawberry
3. Cactus, Pineapple

Actual Answer:

Apple, Apricot, Peach,
Pear, Plum



1. Strawberry
2. Cherry, Strawberry
3. Apple, Strawberry

Actual Answer:

Strawberry



Test Accuracy

1. 13%
2. 15%
3. 4%

Conclusion





99%

Accuracy with SVM for the first question

95%


Accuracy with SVM for 80 Angles

15%

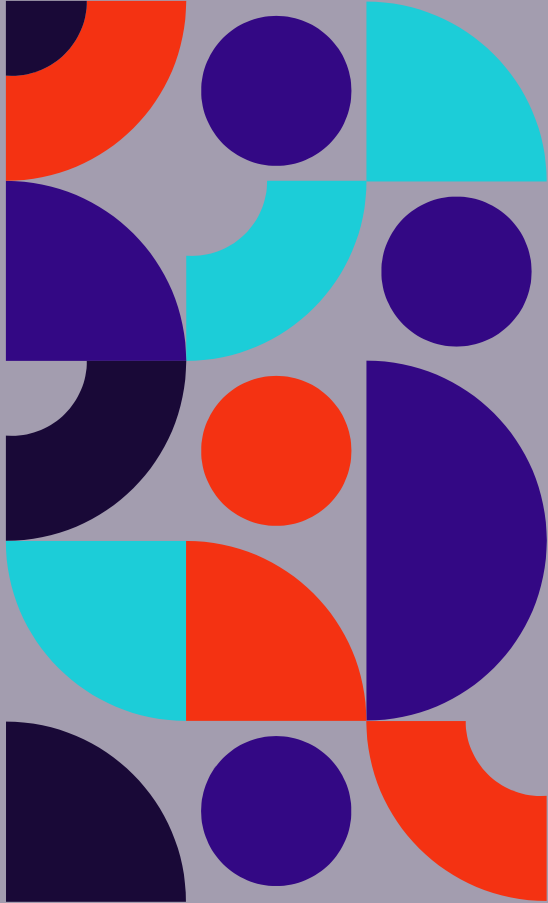
Maximum accuracy for multi label classification with
CNN



Future Work

- 
- Collect multi-label images for a training set for multi-label problem
 - Try other CNN networks such as VGGnet
 - Develop algorithms for fruit grading which can identify the fruit condition/texture





Thanks !

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, infographics & images by Freepik and illustrations by Stories