

Lung Cancer Detection

Sandy Weng



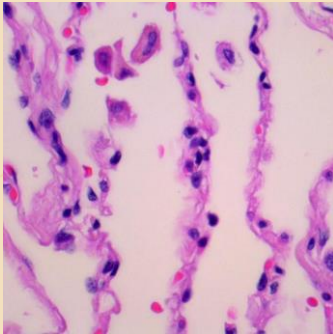


Objective

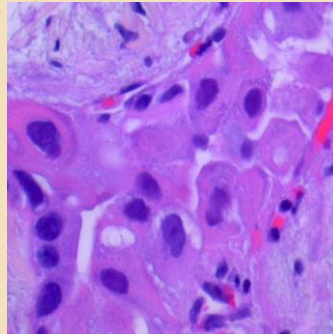
Work with images to create a convolutional neural network to predict whether microscopic lung tissue is cancerous

Data

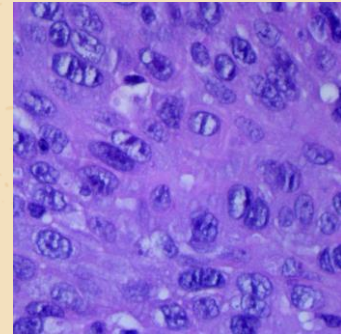
- Kaggle
- 15,000 images
- 3 types of classes



Normal



Lung adenocarcinoma (aca)



Lung squamous cell carcinoma (scc)

Methodology



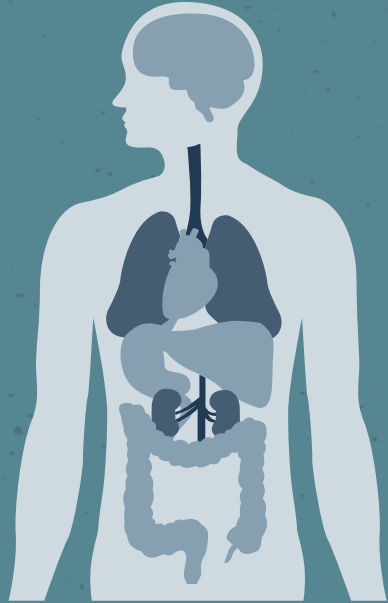
- Tools

- ImageDataGenerator
- Keras
- OpenCV
- Tensorflow



- Models

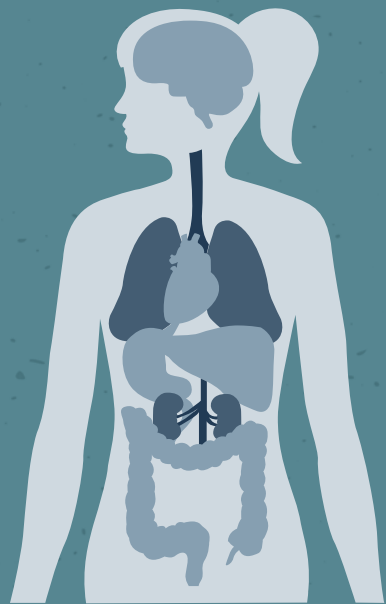
- CNN
- VGG16
- InceptionV3



Baseline Model

KNN

- 400 images
- Classes - Aca and Normal
- Accuracy score: 0.97



Modeling

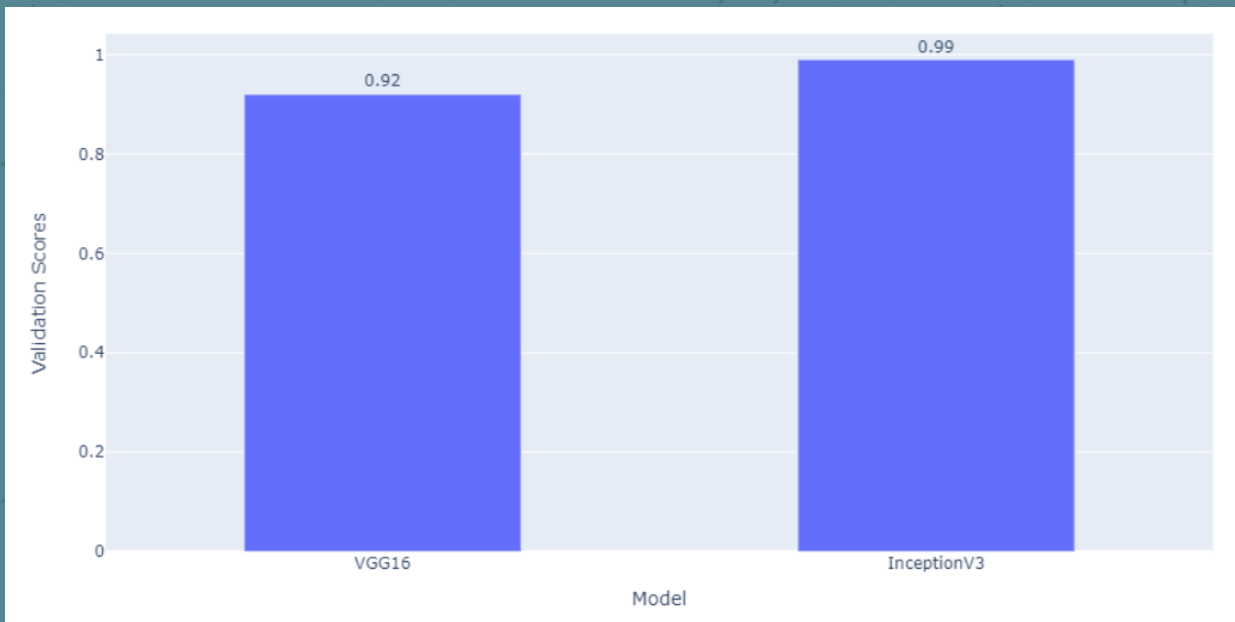
2 Binary Classification Models

- VGG16
- Aca vs Normal
- Aca vs Scc
- Both accuracies ~ 0.97



Modeling

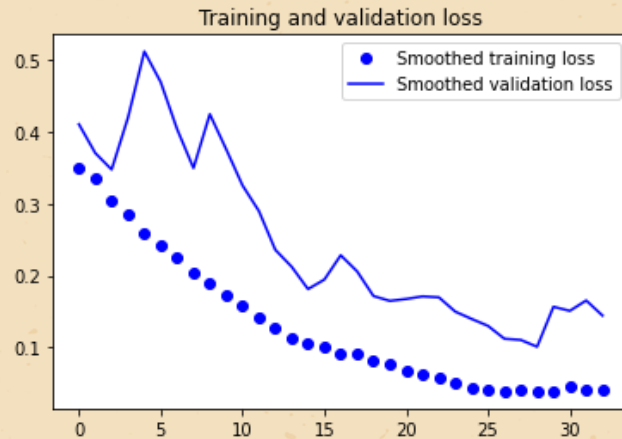
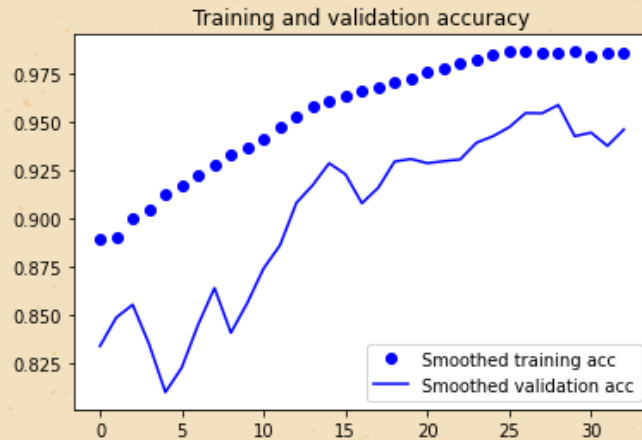
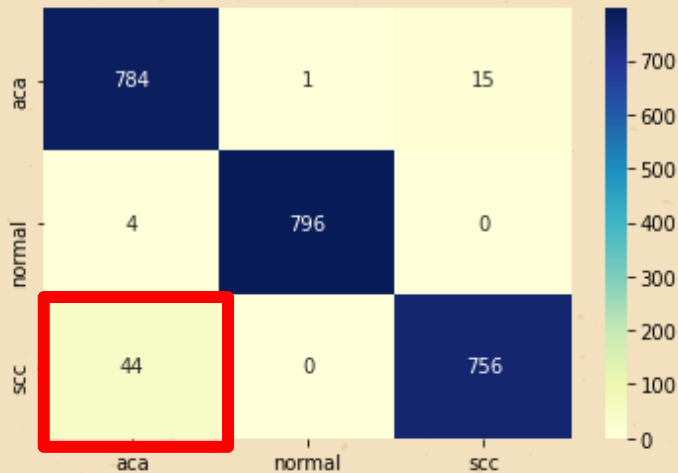
- 3 classes
- 6000 images
- VGG16 validation score: 0.92
- InceptionV3 validation score: 0.99



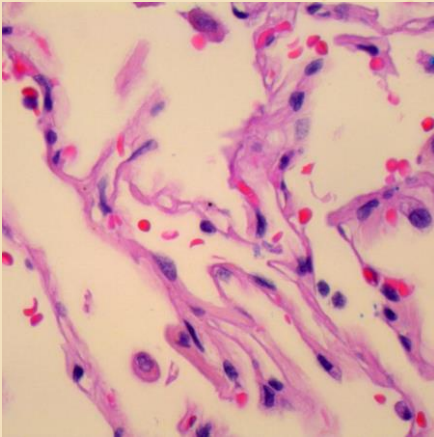
Final Model

InceptionV3

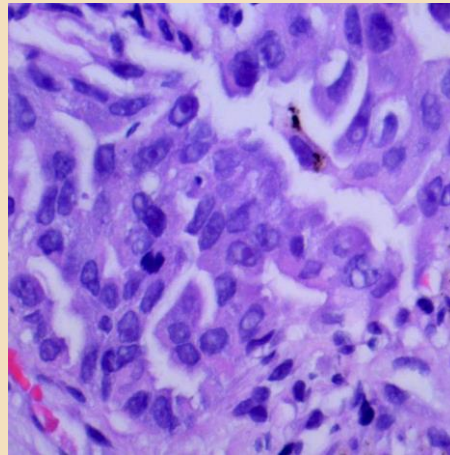
- 3 classes
- 15,000 images
- Accuracy score: 0.973



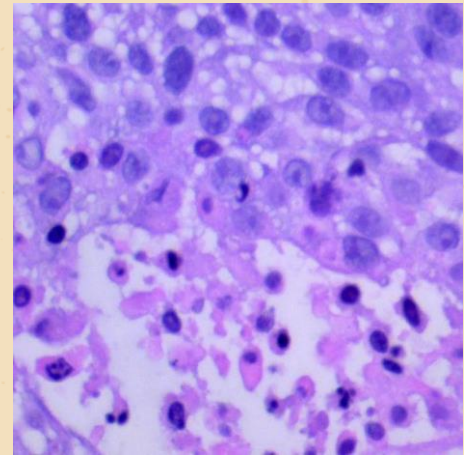
Predictions



Actual: Normal
Predicted: Normal



Actual: Aca
Predicted: Aca



Actual: Scc
Predicted: Scc

Future Work



- Deploy model into Streamlit
- Use Google Cloud to run more models with different parameters to increase accuracy score

THANKS!

Do you have any questions?

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



Appendix



- Final model loss: 0.08
- InceptionV3 small dataset confusion matrix

```
[[438  5  57]
 [  3 497   0]
 [ 34   0 466]]
```

- VGG16 small dataset confusion matrix

```
Confusion Matrix
[[414  24  62]
 [  7 493   0]
 [ 22   0 478]]
```