

Caltech Dataset

Resnet101:

```
data_augmentation = tf.keras.Sequential([
    tf.keras.layers.RandomFlip("horizontal"),
    tf.keras.layers.RandomRotation(0.15),
    tf.keras.layers.RandomZoom(0.2),
    tf.keras.layers.RandomContrast(0.2),
])

base_model = tf.keras.applications.ResNet101(
    input_shape=(IMG_SIZE, IMG_SIZE, 3),
    include_top=False,
    weights="imagenet"
)
base_model.trainable = False # freeze initially

model = tf.keras.Sequential([
    data_augmentation,
    base_model,
    tf.keras.layers.GlobalAveragePooling2D(),
    tf.keras.layers.Dense(256, activation="relu"),
    tf.keras.layers.Dropout(0.4),
    tf.keras.layers.Dense(NUM_CLASSES, activation="softmax")
])

model.summary()
```

I Have done here two type of training.

At first my model was not converging well so i found a technique of **warm up training** . The above architecture is for **warm up training** which will provide you the **trained weights** for further training.

```

base_model.trainable = True
for layer in base_model.layers[:-100]:
    layer.trainable = False

model.compile(
    optimizer=tf.keras.optimizers.Adam(learning_rate=1e-4),
    loss="sparse_categorical_crossentropy",
    metrics=["accuracy"]
)

fine_tune_history = model.fit(
    train_dataset,
    validation_data=val_dataset,
    epochs=20,
    class_weight=class_weights,
    callbacks=callbacks
)

```

Now i have slightly changed the architecture by making model learn weights from data. Now last layers are not using imagenet weights.

Lets see the tweeking of Image Size:(prefferd size=224):

The reason why 224 is preffered because it provide balance in training time and accuracy. Slightly slower than 100 or 128 or 200 but provide more accuracy with 2 to 3% of accuracy margin. If we compare this size to 230++, after fine tuning hyper parameters the bigger image size will slightly improve the accuracy but the training cost will be greater and model train slow.

Now lets see the behavior by seeing the graph of accuracy vs val_accuracy:

Constants:

Batch size = 32

Epochs = 10(warm up) + 20

Desnse layers = 256 with 0.4 dropout

Image Size --> 224

Validation Accuracy (30 epochs): [0.7687819004058838, 0.8118162155151367, 0.8519328832626343, 0.8694383502006531, 0.8708971738815308, 0.8927789926528931, 0.8869438171386719, 0.9022611379623413, 0.8876732587814331, 0.8920496106147766, 0.8832968473434448, 0.9197666049003601, 0.9044492840766907, 0.9321662783622742, 0.9372720718383789, 0.9518599510192871, 0.9438366293907166, 0.9314368963241577, 0.9431072473526001, 0.9613420963287354, 0.9606127142906189, 0.9628008604049683, 0.9584245085716248, 0.9555069208145142, 0.9547775387763977, 0.9606127142906189, 0.964259684085846, 0.9628008604049683, 0.9671772718429565, 0.9686360359191895]

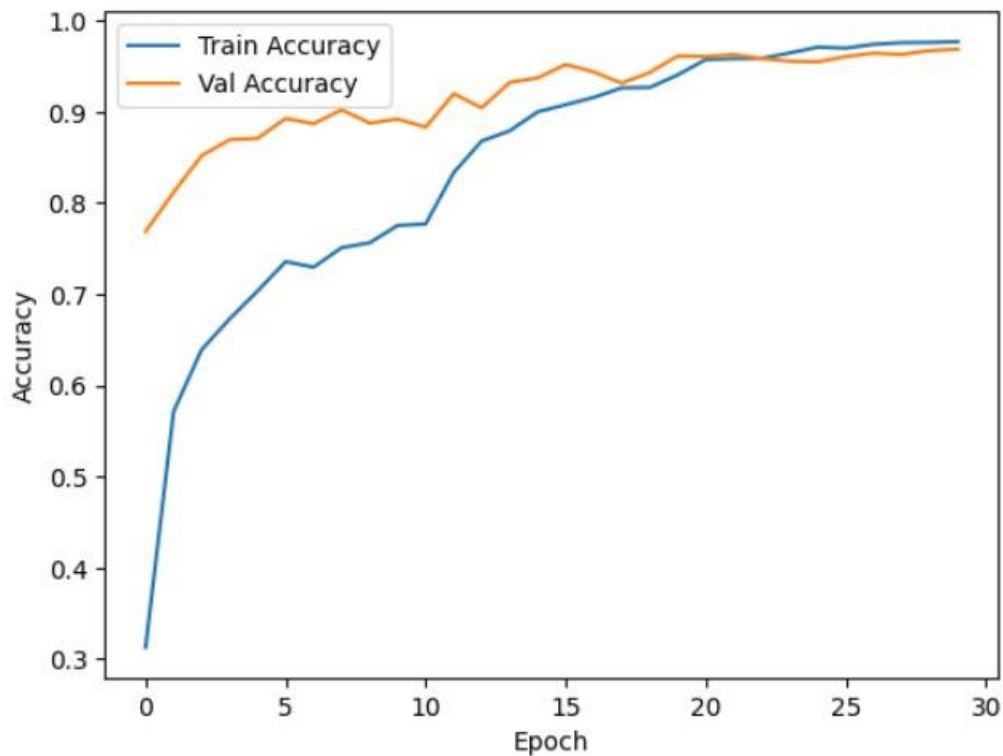


Image size --> 100

Total params: 42,658,176 (162.73 MB)

Trainable params: 0 (0.00 B)

Non-trainable params: 42,658,176 (162.73 MB)

Validation Accuracy (30 epochs): [0.4865061938762665, 0.5426695942878723, 0.5820568799972534, 0.5929977893829346, 0.6039387583732605, 0.6549963355064392, 0.6272793412208557, 0.6819839477539062, 0.6586433053016663, 0.6549963355064392, 0.6418672204017639, 0.7636761665344238, 0.7556527853012085, 0.7760758399963379, 0.7957695126533508, 0.8023340702056885, 0.8059810400009155, 0.8001458644866943, 0.8088986277580261, 0.7731582522392273, 0.8431801795959473, 0.8154631853103638, 0.8132749795913696, 0.8052516579627991, 0.8161925673484802, 0.8351567983627319, 0.8417213559150696, 0.842450737953186, 0.838803768157959, 0.8439095616340637]

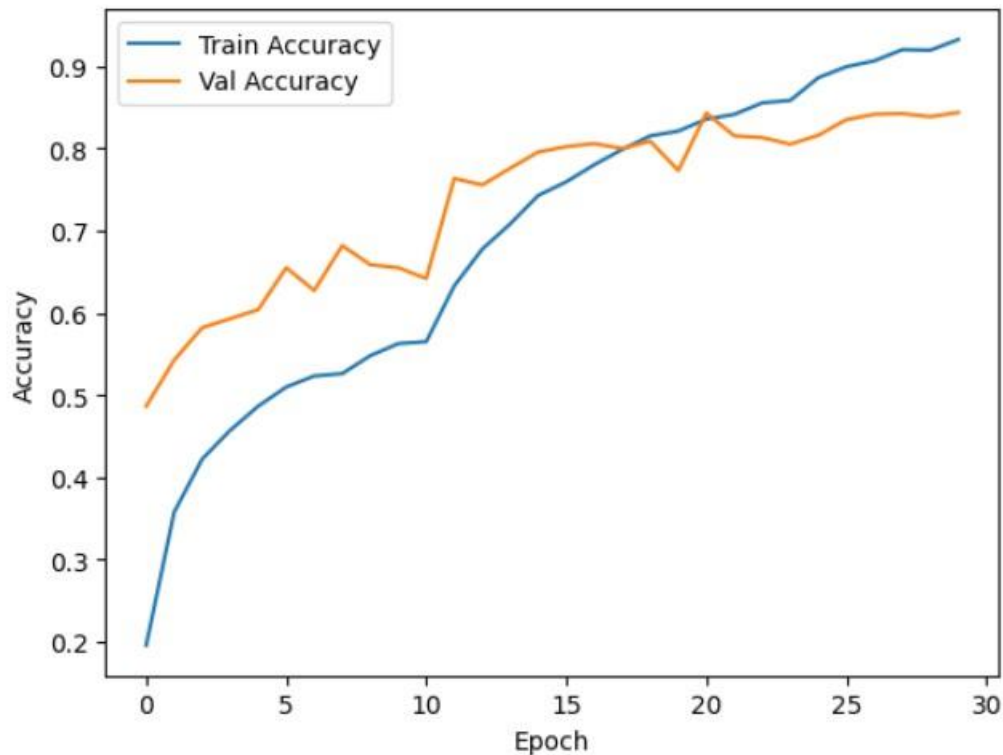


Image size --> 128

Total params: 42,658,176 (162.73 MB)

Trainable params: 0 (0.00 B)

Non-trainable params: 42,658,176 (162.73 MB)

Validation Accuracy (30 epochs): [0.5645514130592346, 0.613420844078064, 0.6761487722396851, 0.7213712334632874, 0.727935791015625, 0.742523729801178, 0.7177242636680603, 0.7819110155105591, 0.7541940212249756, 0.7943107485771179, 0.7760758399963379, 0.8198395371437073, 0.8446389436721802, 0.8563092350959778, 0.8315098285675049, 0.8460977673530579, 0.853391706943512, 0.8687089681625366, 0.8708971738815308, 0.8825674653053284, 0.8920496106147766, 0.8905907869338989, 0.897884726524353, 0.9044492840766907, 0.897884726524353, 0.9008023142814636, 0.9015317559242249, 0.908825695514679, 0.9168490171432495, 0.9226841926574707]

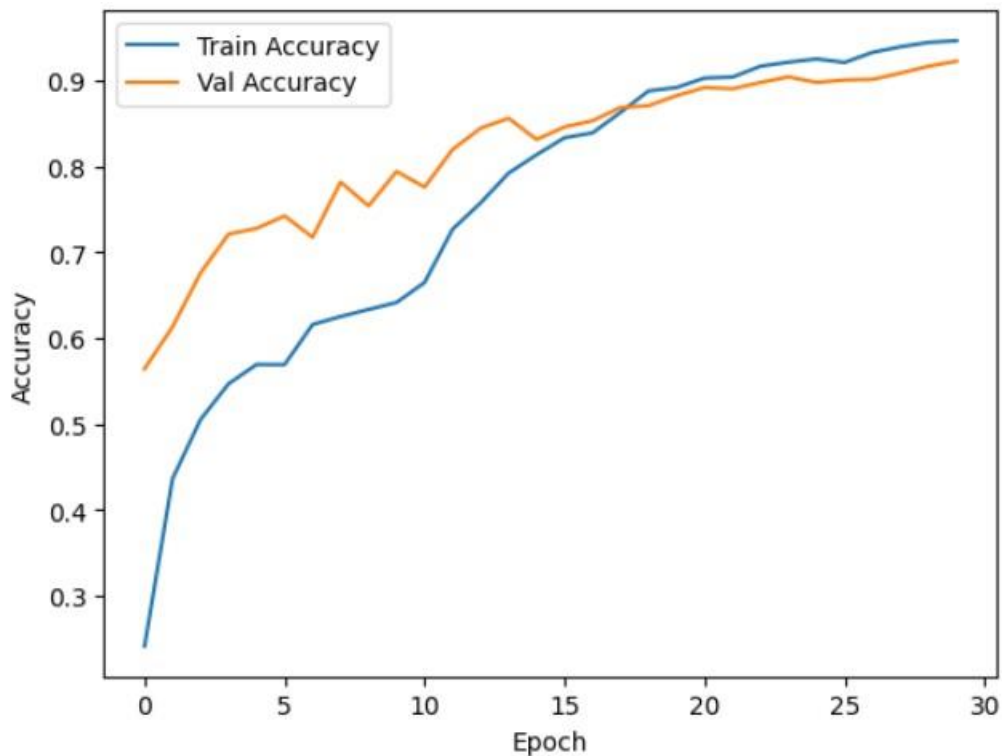


Image size --> 150

Total params: 42,658,176 (162.73 MB)

Trainable params: 0 (0.00 B)

Non-trainable params: 42,658,176 (162.73 MB)

Validation Accuracy (30 epochs): [0.6345732808113098, 0.7133479118347168, 0.7600291967391968, 0.7921225428581238, 0.7913931608200073, 0.8118162155151367, 0.7957695126533508, 0.8344274163246155, 0.8351567983627319, 0.8264040946960449, 0.8344274163246155, 0.8555798530578613, 0.8862144351005554, 0.8913202285766602, 0.8898614048957825, 0.8781911134719849, 0.8781911134719849, 0.9080962538719177, 0.912472665309906, 0.9197666049003601, 0.9153902530670166, 0.9219547510147095, 0.9226841926574707, 0.9241429567337036, 0.9153902530670166, 0.917578399181366, 0.9292487502098083, 0.9219547510147095, 0.9285193085670471, 0.921225368976593]

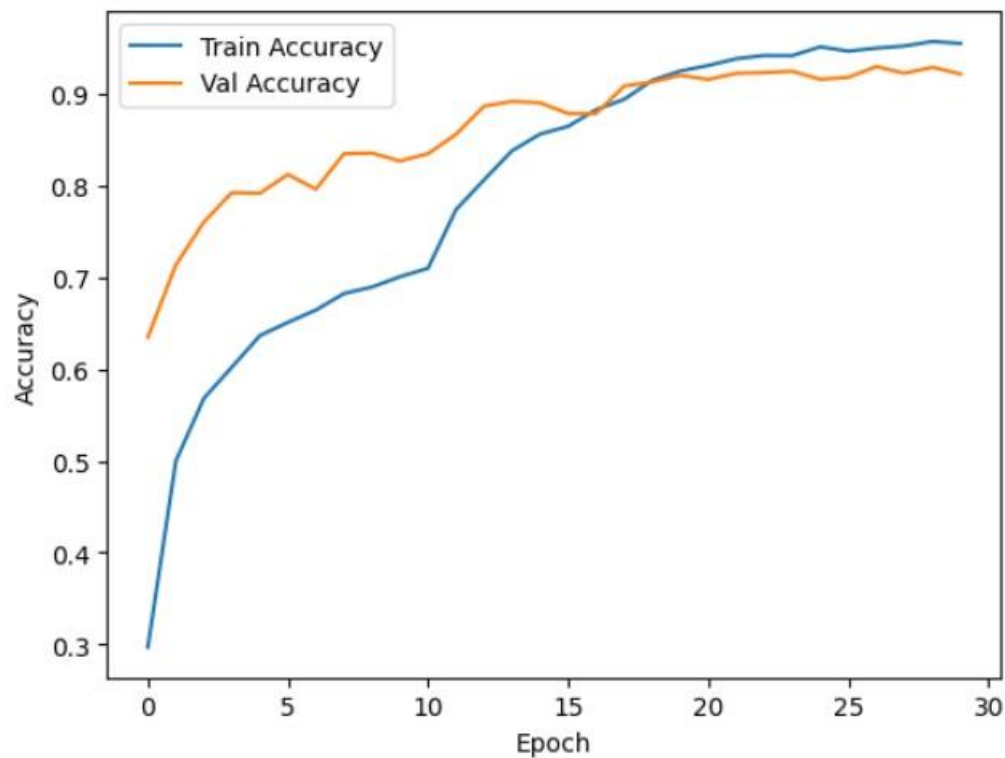


Image size --> 200

Validation Accuracy (30 epochs): [0.7133479118347168, 0.8023340702056885, 0.8154631853103638, 0.8176513314247131, 0.8614150285720825, 0.8460977673530579, 0.8563092350959778, 0.8694383502006531, 0.8643326163291931, 0.8701677322387695, 0.8964259624481201, 0.9037199020385742, 0.9073668718338013, 0.908825695514679, 0.9190372228622437, 0.9190372228622437, 0.908825695514679, 0.9336251020431519, 0.9234135746955872, 0.9241429567337036, 0.9372720718383789, 0.9248723387718201, 0.9256017208099365, 0.9277899265289307, 0.9358132481575012, 0.9321662783622742, 0.939460277557373, 0.9336251020431519, 0.939460277557373, 0.9307075142860413]

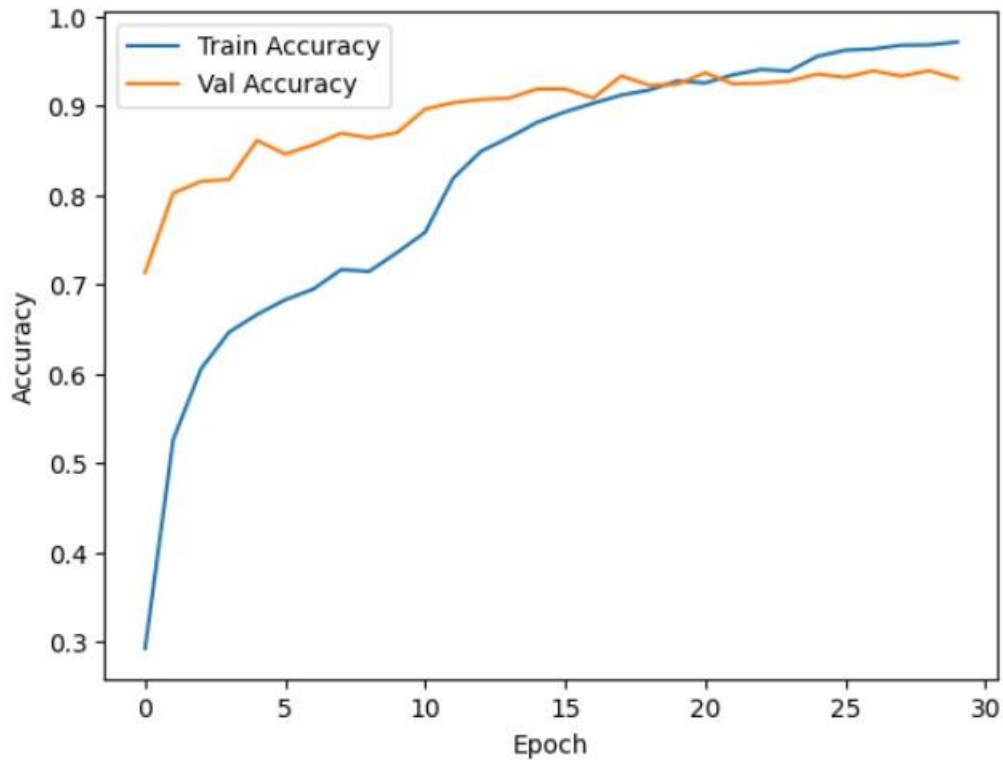


Image size --> 240

Validation Accuracy (30 epochs): [0.7111597657203674, 0.8212983012199402, 0.8490152955055237, 0.8584974408149719, 0.8738147616386414, 0.8628737926483154, 0.8840262293815613, 0.8876732587814331, 0.8636032342910767, 0.894237756729126, 0.8869438171386719, 0.8913202285766602, 0.9029905200004578, 0.912472665309906, 0.9161196351051331, 0.9204959869384766, 0.9350838661193848, 0.9321662783622742, 0.9460247755050659, 0.9161196351051331, 0.9423778057098389, 0.9365426898002625, 0.9533187747001648, 0.9540481567382812, 0.9540481567382812, 0.9540481567382812, 0.9467542171478271, 0.9555069208145142, 0.9540481567382812, 0.9547775387763977]

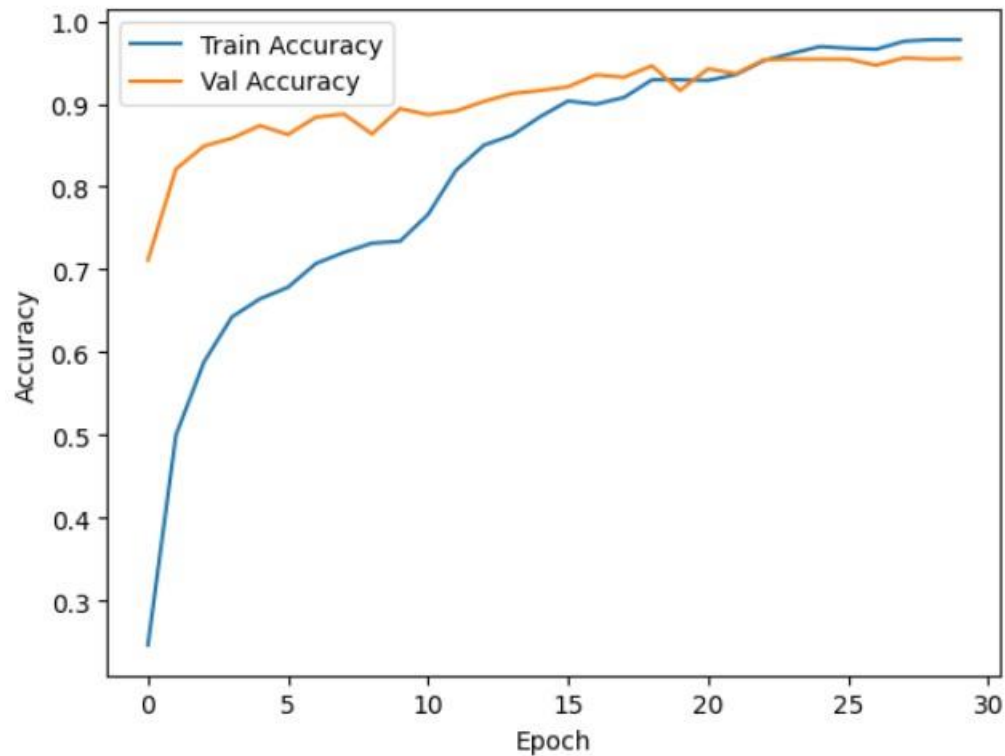
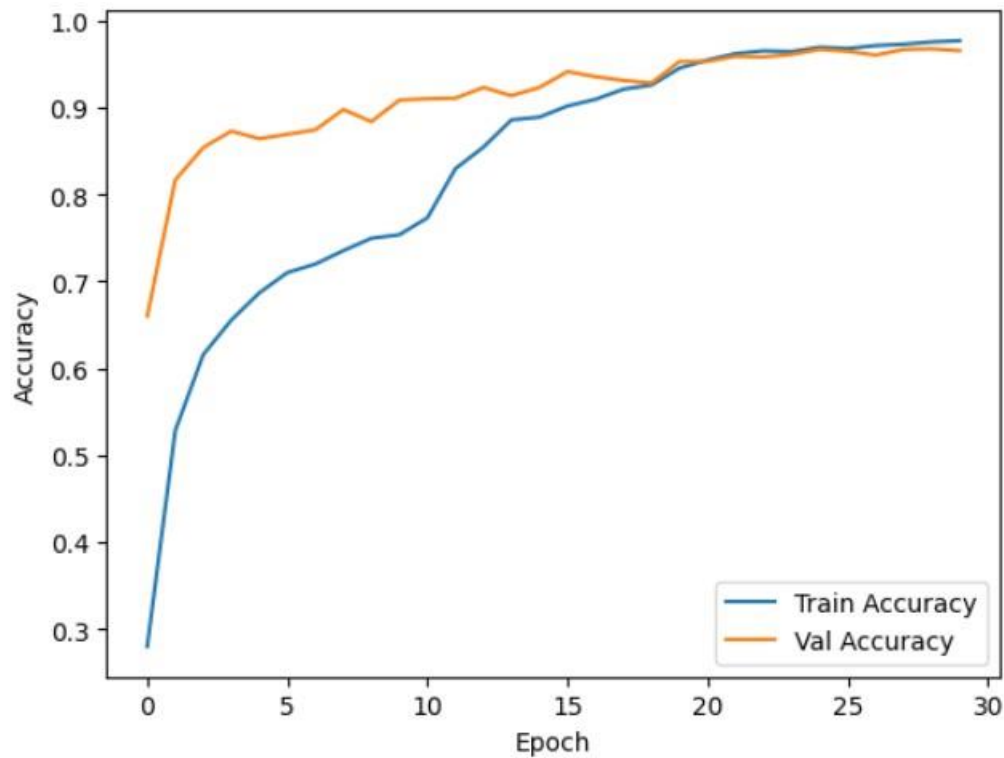
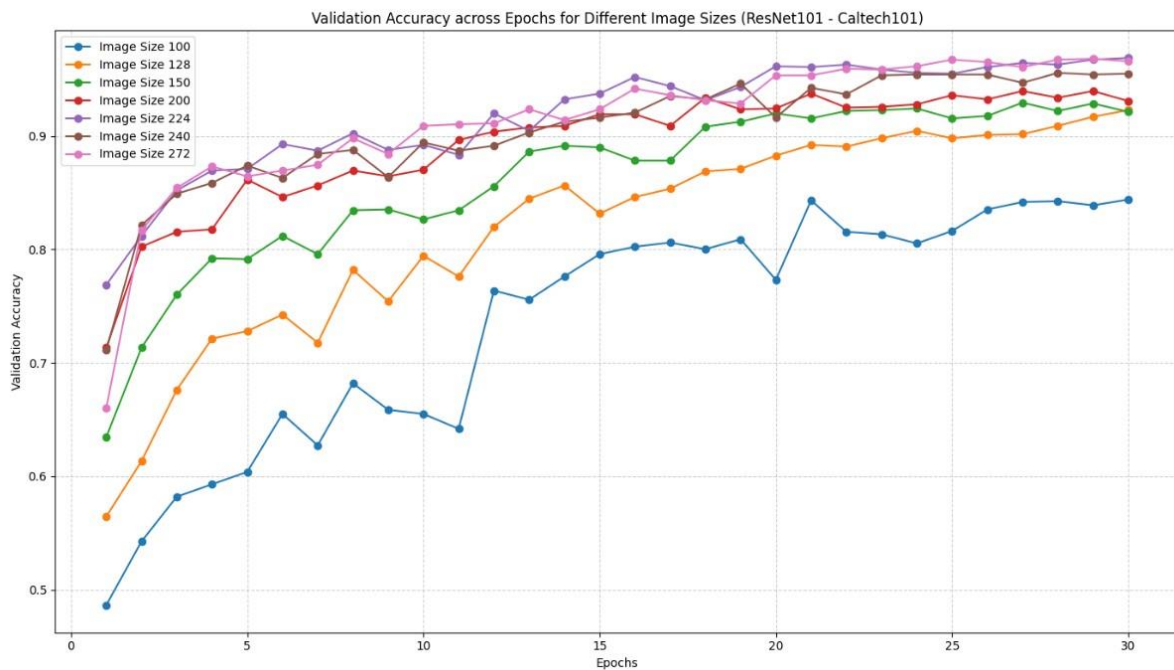


Image size --> 272

Validation Accuracy (30 epochs): [0.660102128982544, 0.8169219493865967, 0.8541210889816284, 0.8730853199958801, 0.8643326163291931, 0.8694383502006531, 0.8745441436767578, 0.897884726524353, 0.8840262293815613, 0.908825695514679, 0.9102844595909119, 0.9110138416290283, 0.9234135746955872, 0.9139314293861389, 0.9234135746955872, 0.9416484236717224, 0.9358132481575012, 0.9314368963241577, 0.9285193085670471, 0.9533187747001648, 0.9533187747001648, 0.9591538906097412, 0.9584245085716248, 0.9613420963287354, 0.9671772718429565, 0.9649890661239624, 0.9606127142906189, 0.9671772718429565, 0.967906653881073, 0.9657184481620789]



Final Comparison



From above comparison the image size 224 is better time wise and accuracy wise in further experiments i will be using 224 image size to see the effect of batch size on training.

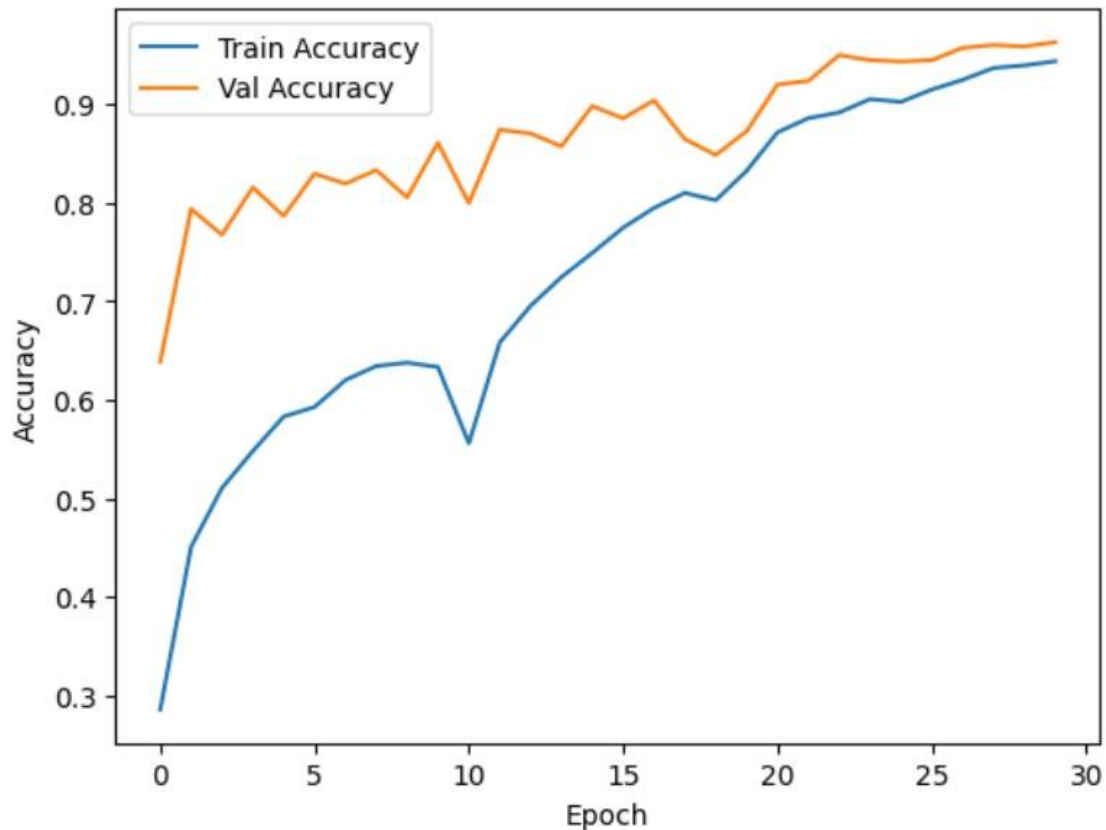
Batch Size Tweaking:

Image size = 224 (for all)

Epochs = 30

Batch size --> 6

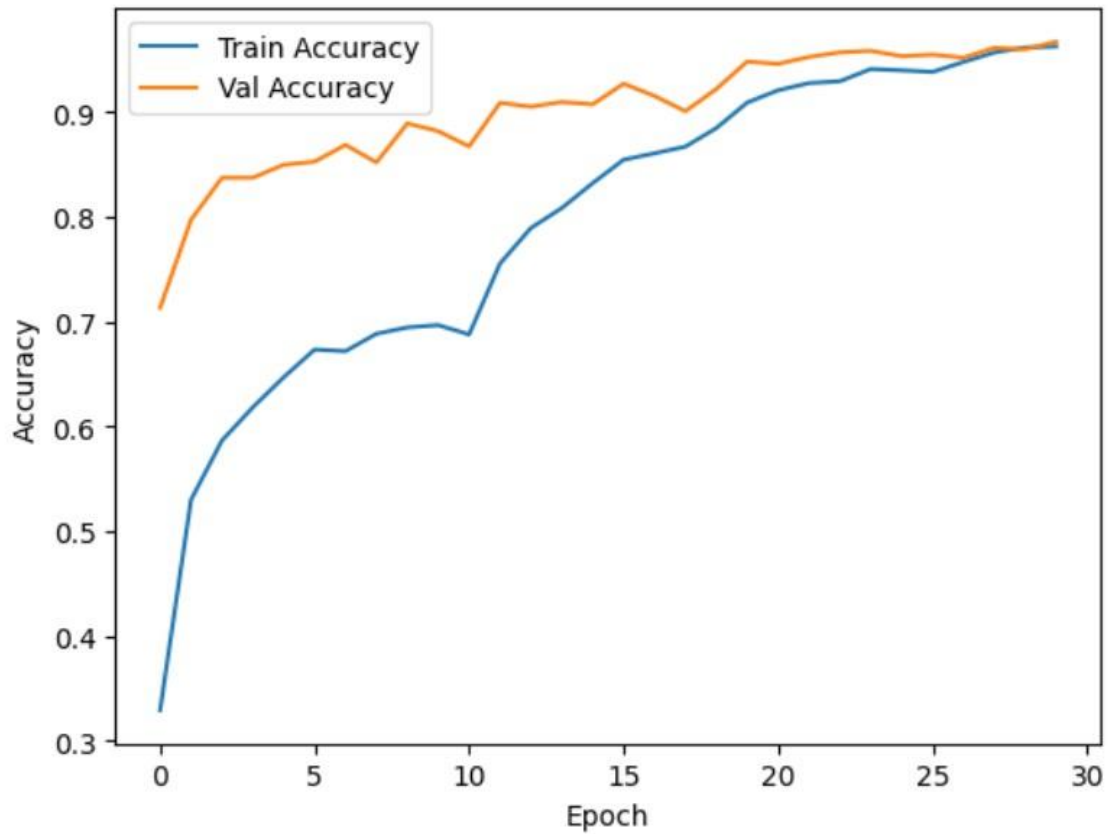
Validation Accuracy (30 epochs): [0.6382202506065369, 0.7935813069343567, 0.7673231363296509, 0.8154631853103638, 0.7862873673439026, 0.8293216824531555, 0.8191101551055908, 0.8329686522483826, 0.8052516579627991, 0.8606856465339661, 0.7994164824485779, 0.8738147616386414, 0.8701677322387695, 0.857038676738739, 0.897884726524353, 0.885485053062439, 0.9037199020385742, 0.8643326163291931, 0.8482859134674072, 0.8723559379577637, 0.9197666049003601, 0.9234135746955872, 0.949671745300293, 0.944566011428833, 0.9431072473526001, 0.944566011428833, 0.9569657444953918, 0.9598832726478577, 0.9584245085716248, 0.9628008604049683]



Bath size --> 12

Final Test Accuracy: 0.9635834097862244

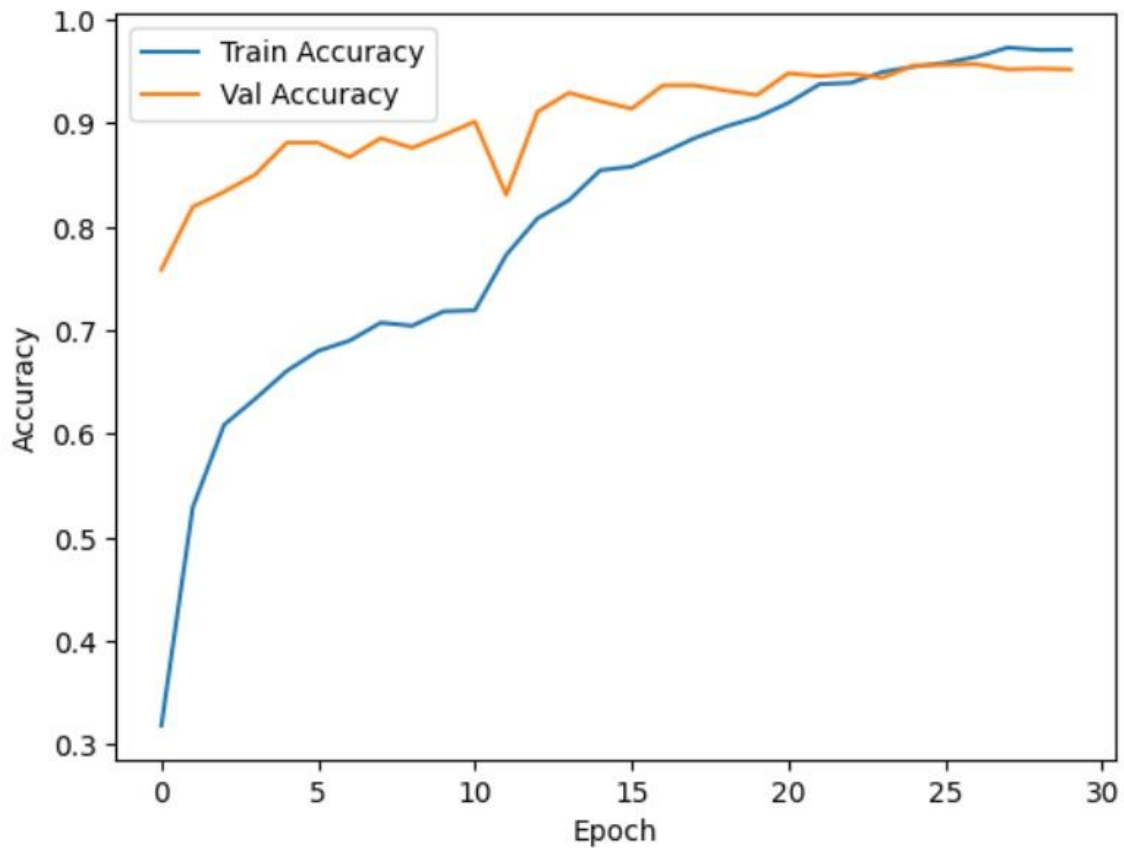
Validation Accuracy (30 epochs): [0.7133479118347168, 0.7972282767295837, 0.8373450040817261, 0.8373450040817261, 0.8497447371482849, 0.8526622653007507, 0.8687089681625366, 0.8519328832626343, 0.889132022857666, 0.8818380832672119, 0.8672502040863037, 0.908825695514679, 0.9051787257194519, 0.9095550775527954, 0.9073668718338013, 0.9270605444908142, 0.9153902530670166, 0.9008023142814636, 0.9219547510147095, 0.9482129812240601, 0.9460247755050659, 0.9525893330574036, 0.9569657444953918, 0.9584245085716248, 0.9533187747001648, 0.9547775387763977, 0.9518599510192871, 0.9613420963287354, 0.9598832726478577, 0.9671772718429565]



Bath size --> 16

Final Test Accuracy: 0.9613984227180481

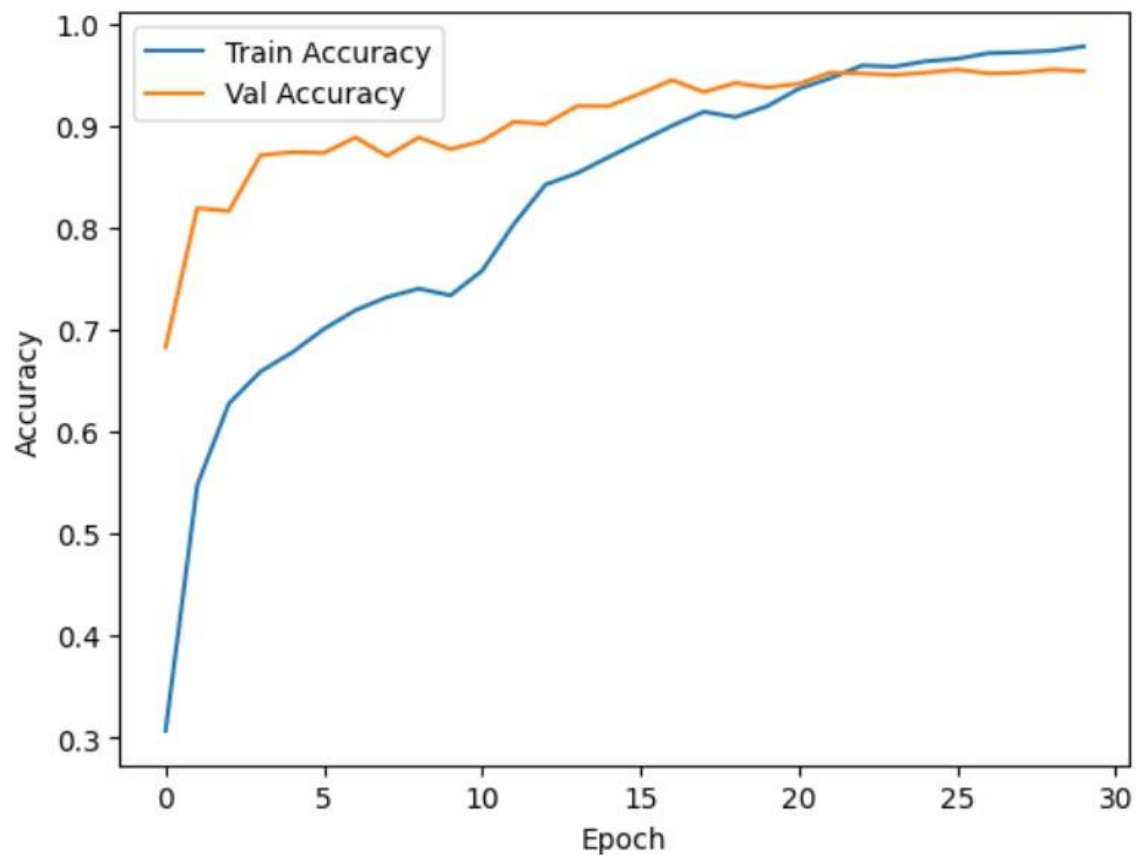
Validation Accuracy (30 epochs): [0.7585703730583191, 0.8191101551055908, 0.833698034286499, 0.8504741191864014, 0.8811087012290955, 0.8811087012290955, 0.8672502040863037, 0.885485053062439, 0.8760029077529907, 0.8884026408195496, 0.9015317559242249, 0.8307804465293884, 0.9110138416290283, 0.9292487502098083, 0.921225368976593, 0.9139314293861389, 0.9365426898002625, 0.9365426898002625, 0.9314368963241577, 0.9270605444908142, 0.9482129812240601, 0.9452953934669495, 0.9474835991859436, 0.9438366293907166, 0.9555069208145142, 0.9562363028526306, 0.9569657444953918, 0.9518599510192871, 0.9525893330574036, 0.9518599510192871]



Bath size --> 24

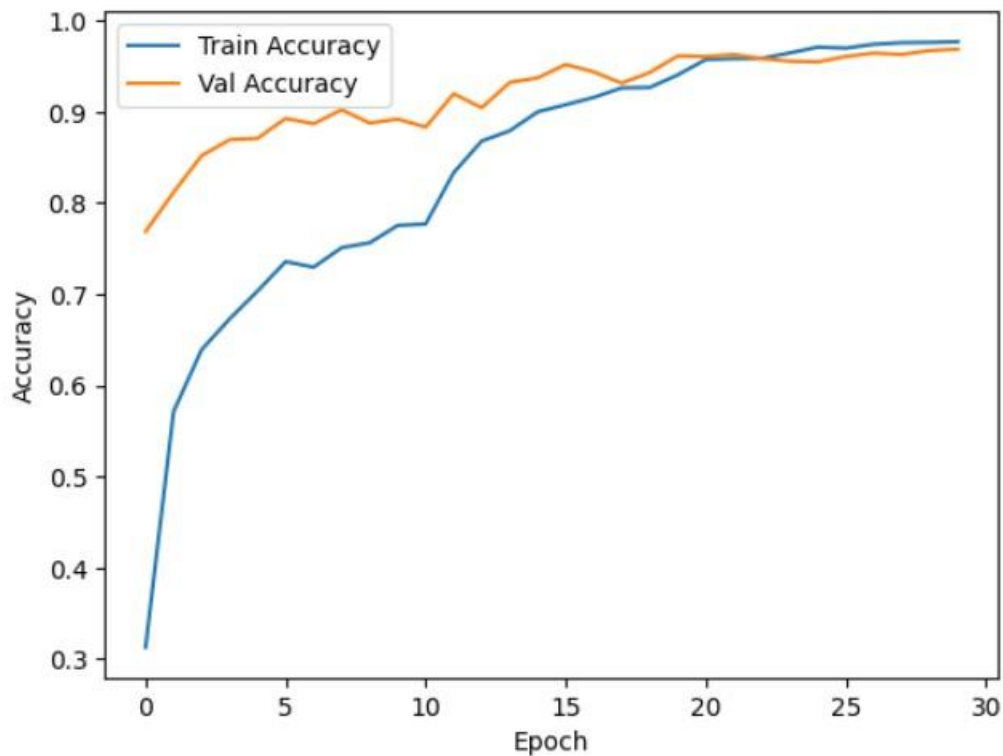
Final Test Accuracy: 0.9584850668907166

Validation Accuracy (30 epochs): [0.6827133297920227, 0.8191101551055908, 0.8161925673484802, 0.8708971738815308, 0.8738147616386414, 0.8730853199958801, 0.8884026408195496, 0.8701677322387695, 0.8884026408195496, 0.8767322897911072, 0.8847556710243225, 0.9037199020385742, 0.9015317559242249, 0.9190372228622437, 0.9190372228622437, 0.9314368963241577, 0.944566011428833, 0.9328957200050354, 0.9416484236717224, 0.9372720718383789, 0.940919041633606, 0.9518599510192871, 0.9511305689811707, 0.949671745300293, 0.9518599510192871, 0.9547775387763977, 0.9511305689811707, 0.9518599510192871, 0.9547775387763977, 0.9533187747001648]



Bath size --> 32

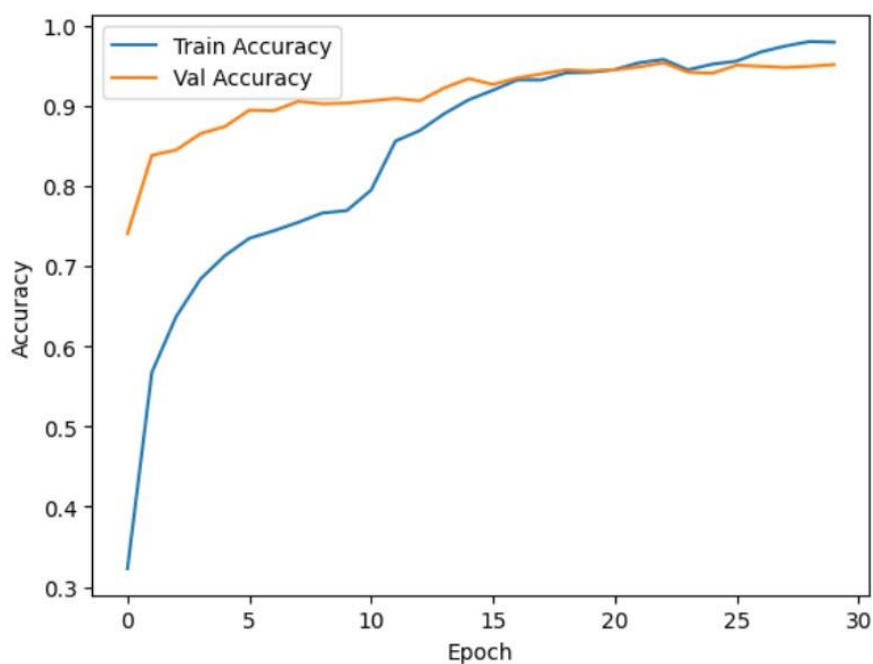
Validation Accuracy (30 epochs): [0.7687819004058838, 0.8118162155151367, 0.8519328832626343, 0.8694383502006531, 0.8708971738815308, 0.8927789926528931, 0.8869438171386719, 0.9022611379623413, 0.8876732587814331, 0.8920496106147766, 0.8832968473434448, 0.9197666049003601, 0.9044492840766907, 0.9321662783622742, 0.9372720718383789, 0.9518599510192871, 0.9438366293907166, 0.9314368963241577, 0.9431072473526001, 0.9613420963287354, 0.9606127142906189, 0.9628008604049683, 0.9584245085716248, 0.9555069208145142, 0.9547775387763977, 0.9606127142906189, 0.964259684085846, 0.9628008604049683, 0.9671772718429565, 0.9686360359191895]



Bath size --> 48

Validation Accuracy 0.9512017369270325

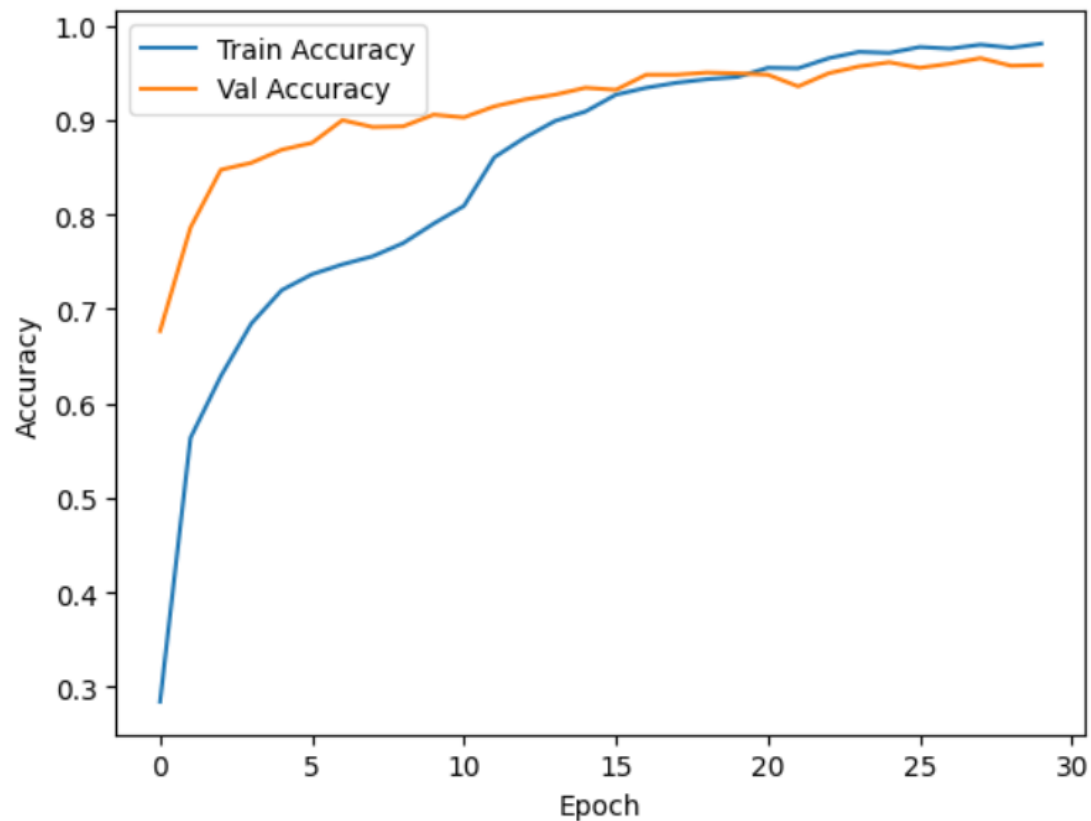
Validation Accuracy (30 epochs): [0.7403355240821838, 0.8380743861198425, 0.8446389436721802, 0.8650619983673096, 0.8738147616386414, 0.894237756729126, 0.8935083746910095, 0.9051787257194519, 0.9022611379623413, 0.9029905200004578, 0.9059081077575684, 0.908825695514679, 0.9059081077575684, 0.9219547510147095, 0.9336251020431519, 0.9263311624526978, 0.9343544840812683, 0.939460277557373, 0.944566011428833, 0.9431072473526001, 0.944566011428833, 0.9482129812240601, 0.9533187747001648, 0.9416484236717224, 0.9401896595954895, 0.9504011869430542, 0.9489423632621765, 0.9474835991859436, 0.9489423632621765, 0.9511305689811707]



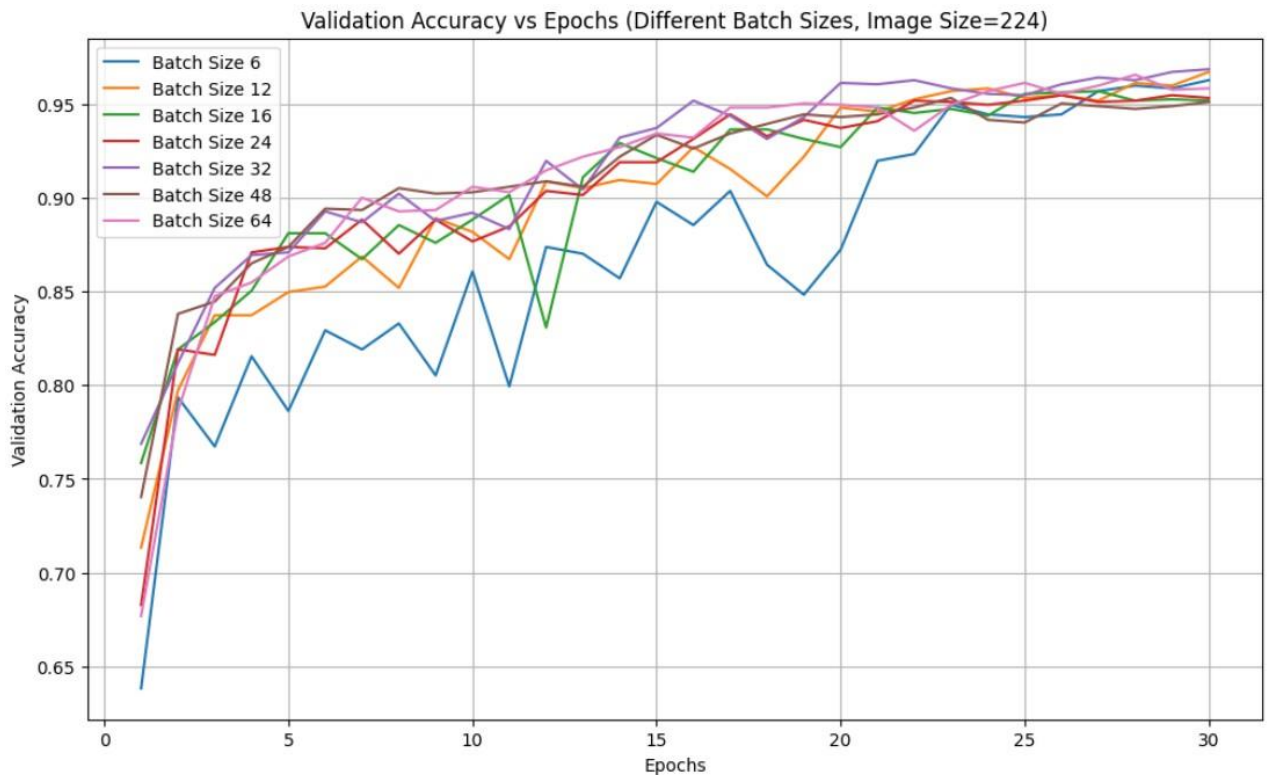
Bath size --> 64

Validation Accuracy 0.9694100618362427

Validation Accuracy (30 epochs): [0.6768782138824463, 0.7862873673439026, 0.8475565314292908, 0.8548504710197449, 0.8687089681625366, 0.8760029077529907, 0.9000729322433472, 0.8927789926528931, 0.8935083746910095, 0.9059081077575684, 0.9029905200004578, 0.9146608114242554, 0.9219547510147095, 0.9270605444908142, 0.9343544840812683, 0.9321662783622742, 0.9482129812240601, 0.9482129812240601, 0.9504011869430542, 0.949671745300293, 0.9482129812240601, 0.9358132481575012, 0.949671745300293, 0.9569657444953918, 0.9613420963287354, 0.9555069208145142, 0.9598832726478577, 0.9657184481620789, 0.9576951265335083, 0.9584245085716248]



Overall comparison batchwise:



Tunning The Dense layers:

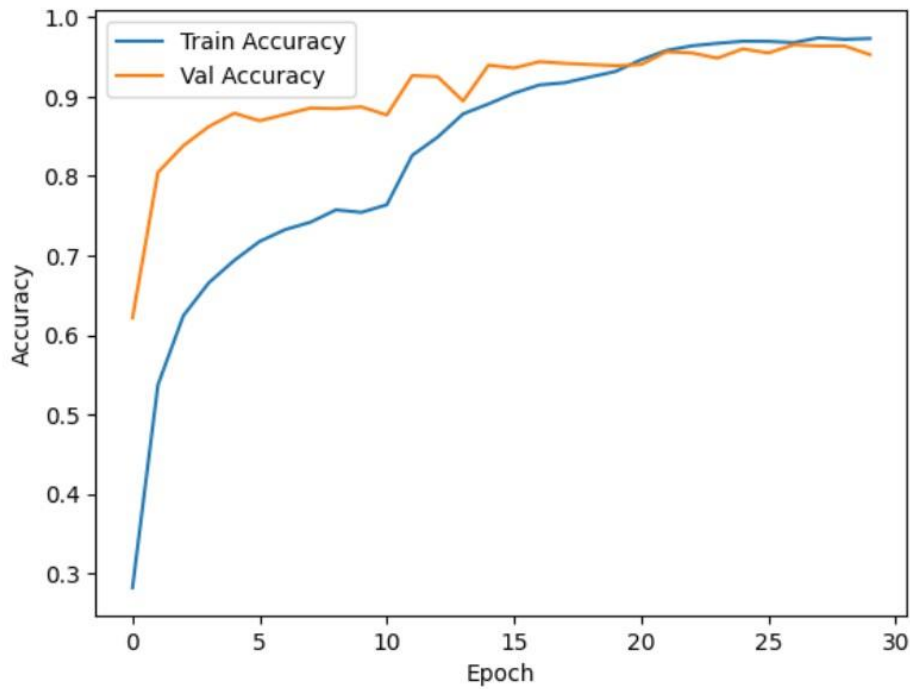
Constants:

- Batch size 32(according to available resources)
- Epoch 30
- Image size 224
- Resnet last 100 trainable

Single layer 128 (Dropout = 0.3 , Activation = Relu):

Final Test Accuracy: 0.959213376045227

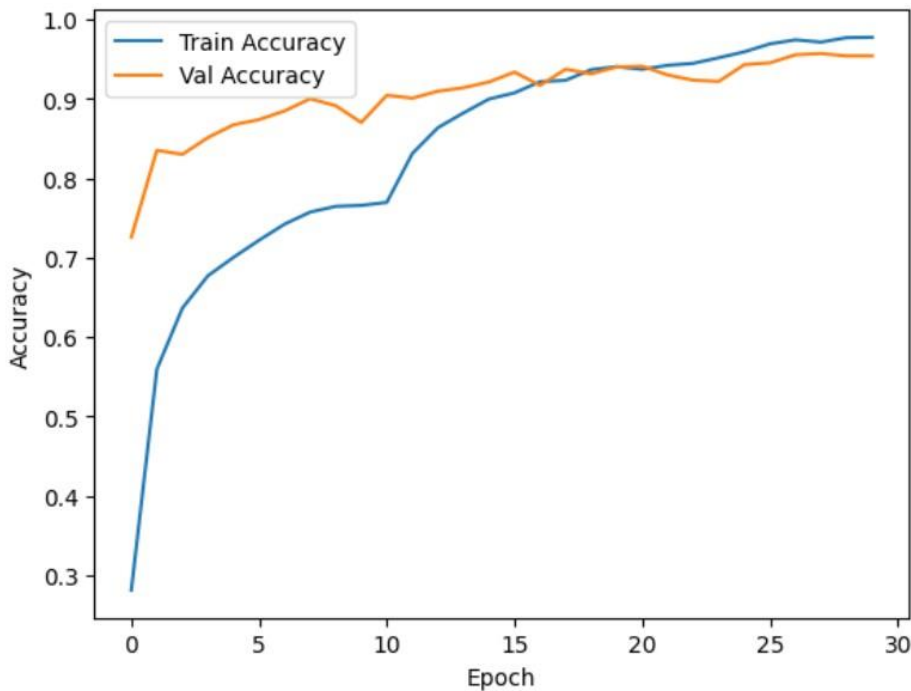
Validation Accuracy (30 epochs): [0.6214442253112793, 0.8045222759246826, 0.8380743861198425, 0.862144410610199, 0.8789204955101013, 0.8694383502006531, 0.8774617314338684, 0.885485053062439, 0.8847556710243225, 0.8869438171386719, 0.8767322897911072, 0.9263311624526978, 0.9248723387718201, 0.894237756729126, 0.939460277557373, 0.9358132481575012, 0.9438366293907166, 0.9416484236717224, 0.9401896595954895, 0.9387308359146118, 0.9401896595954895, 0.9562363028526306, 0.9547775387763977, 0.9482129812240601, 0.9598832726478577, 0.9547775387763977, 0.9649890661239624, 0.9635302424430847, 0.9635302424430847, 0.9525893330574036]



Single layer 128 (Dropout = 0.3 , Activation = mish):

Final Test Accuracy: 0.959213376045227

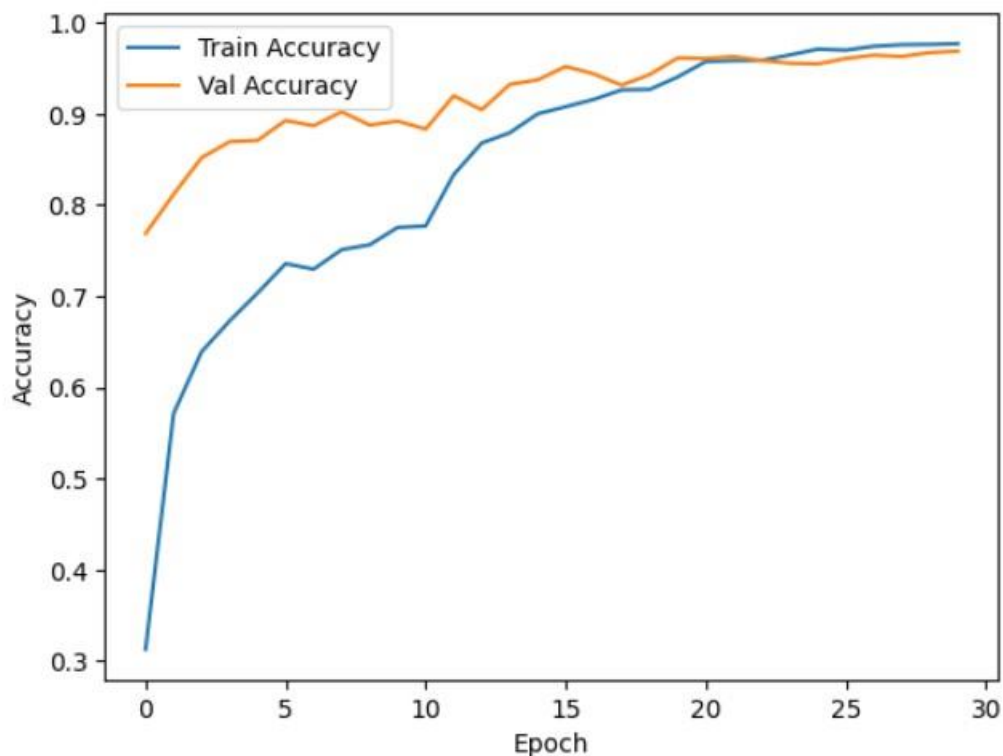
Validation Accuracy (30 epochs): [0.7257476449012756, 0.8351567983627319, 0.830051064491272, 0.8512035012245178, 0.8672502040863037, 0.8738147616386414, 0.8847556710243225, 0.9000729322433472, 0.8913202285766602, 0.8701677322387695, 0.9044492840766907, 0.9008023142814636, 0.9095550775527954, 0.9139314293861389, 0.921225368976593, 0.9336251020431519, 0.9168490171432495, 0.9372720718383789, 0.9314368963241577, 0.9401896595954895, 0.940919041633606, 0.9299781322479248, 0.9234135746955872, 0.9219547510147095, 0.9431072473526001, 0.9452953934669495, 0.9555069208145142, 0.9569657444953918, 0.9540481567382812, 0.9540481567382812]



Single layer 256 (Dropout = 0.4 , Activation = relu):

Final Test Accuracy: 0.964259684085846

Validation Accuracy (30 epochs): [0.7687819004058838, 0.8118162155151367, 0.8519328832626343, 0.8694383502006531, 0.8708971738815308, 0.8927789926528931, 0.8869438171386719, 0.9022611379623413, 0.8876732587814331, 0.8920496106147766, 0.8832968473434448, 0.9197666049003601, 0.9044492840766907, 0.9321662783622742, 0.9372720718383789, 0.9518599510192871, 0.9438366293907166, 0.9314368963241577, 0.9431072473526001, 0.9613420963287354, 0.9606127142906189, 0.9628008604049683, 0.9584245085716248, 0.9555069208145142, 0.9547775387763977, 0.9606127142906189, 0.964259684085846, 0.9628008604049683, 0.9671772718429565, 0.9686360359191895]



Single layer 256 (Dropout = 0.4 , Activation = mish):

Final Test Accuracy: 0.9533867239952087

Validation Accuracy (30 epochs): [0.7447118759155273, 0.8380743861198425, 0.8490152955055237, 0.8738147616386414, 0.885485053062439, 0.8898614048957825, 0.897884726524353, 0.9073668718338013, 0.9022611379623413, 0.917578399181366, 0.8920496106147766, 0.9183077812194824, 0.9139314293861389, 0.9204959869384766, 0.9256017208099365, 0.9423778057098389, 0.939460277557373, 0.9387308359146118, 0.9358132481575012, 0.9438366293907166, 0.9504011869430542, 0.9467542171478271, 0.9489423632621765, 0.9460247755050659, 0.9533187747001648, 0.9547775387763977, 0.9555069208145142, 0.9547775387763977, 0.9569657444953918, 0.9504011869430542]

