Transfer Learning On Bollywood actor Dataset With Mobilenet Large and Small Weights.

In this file we will explore the use and effectiveness of transfer learning using both small and large mobilenet v3 on a customized dataset of 12 bollywood actor faces taken from (https://www.kaggle.com/datasets/sushilyadav1998/bollywood-celeb-localized-face-dataset).

Taking a closer look at dataset, we see that it has 12 classes (one for each of the 12 actors) and aroud 1137 training images.

The dataset is large enough to train smaller a model from scratch but only with many epochs (and likely overfitting the data since there are so few train images). In the case where we don't have large enough data, using transfer learning can be a great option. Let's see how well mobilenet v3 model performs on this dataset. We will experiment trying to use 5, 10, 20, 30, and 40 training images for each class (amounting to a total of 60, 120, 240, 360, and 480 training images) to leaving 657 images for testing.

Note that this was run in the anaconda environment environment.yml included in the submission of this assignment.

Note: IDK why the text in the cell above isnt rendering correctly... inspect a cell if it is not making sense as the text there is correct.

We begin by loading the necessary imports

In [1]:
```
import torch
import torch.nn as nn
import torch.nn.functional as F
from torchvision.models import mobilenet_v3_large, MobileNet_V3_Large_Weights
from torchvision.models import mobilenet_v3_small, MobileNet_V3_Small_Weights
from torchvision import transforms
from torch.utils.data import Dataset, DataLoader
import numpy as np
from torchvision.io import read_image
import os
import matplotlib.pyplot as plt
```

Enable GPU usage

In [2]:
```
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
```

Since we will be using multiple versions of mobilenet (large, small, and pretrained vs not pretrained - for comparison and analysis) we implement a functoin to load the correct model and return it.

In [3]:
```
# https://pytorch.org/vision/stable/models/mobilenetv3.html
# https://pytorch.org/blog/torchvision-mobilenet-v3-implementation/
def get_cleared_model(num_classes = 12, model_type = "large"):
    """
    returns the mobilenet v3 model of type model_type (one of large, small, or small_blank), and modifies it so that the final output layer
    has size num_classes.
    Returns the selected model or nothing no valid model is selected
    """
    if model_type == "large":
        large_model = mobilenet_v3_large(weights=MobileNet_V3_Large_Weights.DEFAULT)
        num_features = large_model.classifier[-1].in_features
        large_model.classifier[-1] = nn.Linear(num_features, num_classes)
        return large_model
    elif model_type == "small":
        small_model = mobilenet_v3_small(weights=MobileNet_V3_Small_Weights.DEFAULT)
        num_features = small_model.classifier[-1].in_features
        small_model.classifier[-1] = nn.Linear(num_features, num_classes)
        return small_model
    elif model_type == "small_blank":
        small_model = mobilenet_v3_small()
        num_features = small_model.classifier[0].in_features
        small_model.classifier = nn.Linear(num_features, num_classes)
        return small_model

    return None
```

Next we will establish the data transform to apply to our images to have them fit the input of the mobile net v3 models

In [4]:
```
transform = transforms.Compose([
    transforms.ToPILImage(mode='RGB'),             # Convert to PIL Image
    transforms.Resize((224, 224)),      # Resize to 224x224
    transforms.ToTensor()                # Convert to PyTorch tensor
])
```

Now we create a function to load the data that makes the correct train/test split to ensure that reguardless of the amount of data in the train section, the test data will remain the same for consistency of analysis. We make this function since we will load a new train/test split many times to see how the amount of training data impacts the performance of the transfer learning.

In [5]:
```
def load_dataset(dataset_path='data/Bollywood_celeb_dataset', training_amount = 20, max_training_amount = 40):
    """
    load_dataset returns X_train, X_test, y_train, y_test of the Bollywood2 dataset found at the path dataset_path
    The dataset must follow a specific form that can be viewed in this directory, if it is deviated from changes may have to be made.
    Since we create our own train/test split we expect there to be one set of labels and one file of images that correspond.
    training_amount is the amount of data we will want to be in the training set
    max_training_amount is the max amount of training data we will use at all in the program (necessary to ensure test dataset is the same for all tra
```

```
    """
    # initializations
    X_train = []
    X_test = []
    label_train = []
    label_test = []
    name_to_label = {}
    label_count = 0

    # iterate over each person (folder) in the dataset
    for person_folder in os.listdir(dataset_path):
        person_path = os.path.join(dataset_path, person_folder)
        if os.path.isdir(person_path):
            name_to_label[person_folder] = label_count # set the label for this person
            person_images = []

            # iterate over all images for the person and load them
            for image_file in os.listdir(person_path):
                image_path = os.path.join(person_path, image_file)
                im = read_image(image_path)
                if im is not None:
                    person_images.append(im)

            # create the train/test split for the person
            face_train, face_test, y_train, y_test = split_data(person_images, np.full(len(person_images), label_count), training_amount, max_training

            label_count += 1

            # add the person to the overall train/test data split
            X_train = X_train + face_train
            X_test = X_test + face_test
            label_train = label_train + y_train
            label_test = label_test + y_test

    X_train = np.array(X_train)
    X_test = np.array(X_test)
    label_train = np.array(label_train)
    label_test = np.array(label_test)

    return X_train.transpose((0, 2, 3, 1)), X_test.transpose((0, 2, 3, 1)), label_train, label_test

def split_data(X, y, training_amount = 20, max_training_amount = 40):
    """
    function to split the data (X, y) into train/test split where each actor has training_amount training images
    and number of images - max_training_amount of test images
    """
    training_amount = [i for i in range(min(len(y), training_amount))]
    testing_amount = [i for i in range(len(y) - max_training_amount)]

    return [X[i] for i in training_amount], [X[i] for i in testing_amount], [y[i] for i in training_amount], [y[i] for i in testing_amount]
```

Now we will create a class that can be used as a dataset in a pytorch dataloader. For this we must implemnet init, len, and getitem

In [6]:
```
class BollywoodCelebDataset(Dataset):
    def __init__(self, X, label, transform=None):
        """
        X is the images, label is the labels, transform is a transform to apply on them if desired.
        """
        self.X = X
        self.label = label
        self.transform = transform

    def __len__(self):
        return len(self.X)

    def __getitem__(self, idx):
        """
        return the idx item of the dataset
        """
        image = self.X[idx]
        label = self.label[idx]

        if self.transform:
            image = self.transform(image)

        return image, label
```

Now we create a train function to run one epoch of training, which allows us more flexibility to store results as desired later.

In [7]:
```
def train(model, train_loader, criterion = nn.CrossEntropyLoss()):
    """
    trains model on train_loader using criterion as the loss function and Adam optimizer
    returns the loss and accuracy of the epoch
    """
    # initializations
    optimizer = torch.optim.Adam(model.parameters(), lr=0.001)
    size = len(train_loader.dataset)
    num_batches = len(train_loader)

    # train the model
    model.train()
    total_loss = 0.0
```

```
    correct = 0.0
    for batch, (inputs, labels) in enumerate(train_loader):
        inputs, labels = inputs.to(device), labels.to(device)
        optimizer.zero_grad()
        outputs = model(inputs)
        labels = labels.to(torch.long)
        loss = criterion(outputs, labels)
        loss.backward()
        optimizer.step()

        # loss and accuracy tracking
        total_loss += loss.item()
        correct += (outputs.argmax(1) == labels).type(torch.float).sum().item()

    return total_loss/num_batches, correct/size
```

Create a test function

In [8]:
```
def test(model, test_loader, criterion = nn.CrossEntropyLoss()):
    """
    tests model on test_loader using criterion loss
    returns the loss and accuracy
    """
    # initializations
    size = len(test_loader.dataset)
    num_batches = len(test_loader)
    model.eval()
    total_loss, correct = 0.0, 0.0

    # test
    with torch.no_grad():  # for testing not training
        for X, y in test_loader:
            X, y = X.to(device), y.to(device)
            pred = model(X)
            y = y.to(torch.long)
            total_loss += criterion(pred, y).item()
            correct += (pred.argmax(1) == y).type(torch.float).sum().item()

    return total_loss/num_batches, correct/size
```

Write a function to run a train test sequence that we can apply to each of the model types

In [9]:
```
def train_test_sequence(model_name = "large", save_base_name = None, train_size = [5, 10, 20, 30, 40], epochs = 10, num_classes = 12, dataset_path = '
    """
    model_name -> the model name provided to get_cleared_model. One of large, small, small_blank
    save_base_name -> path to save the model state in
    train_size -> a list of the size of the training set for each class we will run a train/test sequence on
    epochs -> the number of epochs used for each set
    num_classes -> the number of classes in the dataset being used (102 for Oxford Flowers)
    dataset_path -> the path to the Oxford 102 data

    Returns: a 2D list where each row is the train accuracy for each entry of train_size and the column is the performance at that epoch
            and another 2D list of the same shape for the test accuracy
    """
    train_size_train_accuracy = []
    train_size_test_accuracy = []

    # for each size of the training set:
    for i in train_size:
        epoch_train_accuracy = []
        epoch_test_accuracy = []

        # data setup
        X_train, X_test, y_train, y_test = load_dataset(dataset_path, i, max(train_size))
        train_ds = BollywoodCelebDataset(X_train, y_train, transform)
        train_loader = DataLoader(train_ds, batch_size=32, shuffle=True)
        test_ds = BollywoodCelebDataset(X_test, y_test, transform)
        test_loader = DataLoader(test_ds, batch_size=32, shuffle=True)

        # model setup
        checkpoint = None
        if save_base_name:
            try:
                # Attempt to load the checkpoint
                checkpoint = torch.load(f'{save_base_name}_{i}.pth')
            except FileNotFoundError:
                # If the file doesn't exist, set checkpoint to None
                checkpoint = None
        model = get_cleared_model(num_classes, model_name)
        if checkpoint:
            model.load_state_dict(checkpoint)
        model = model.to(device)

        # train and test
        for epoch in range(epochs):
            train_loss, train_accuracy = train(model, train_loader)

            test_loss, test_accuracy = test(model, test_loader)

            epoch_train_accuracy.append(train_accuracy)
            epoch_test_accuracy.append(test_accuracy)
```

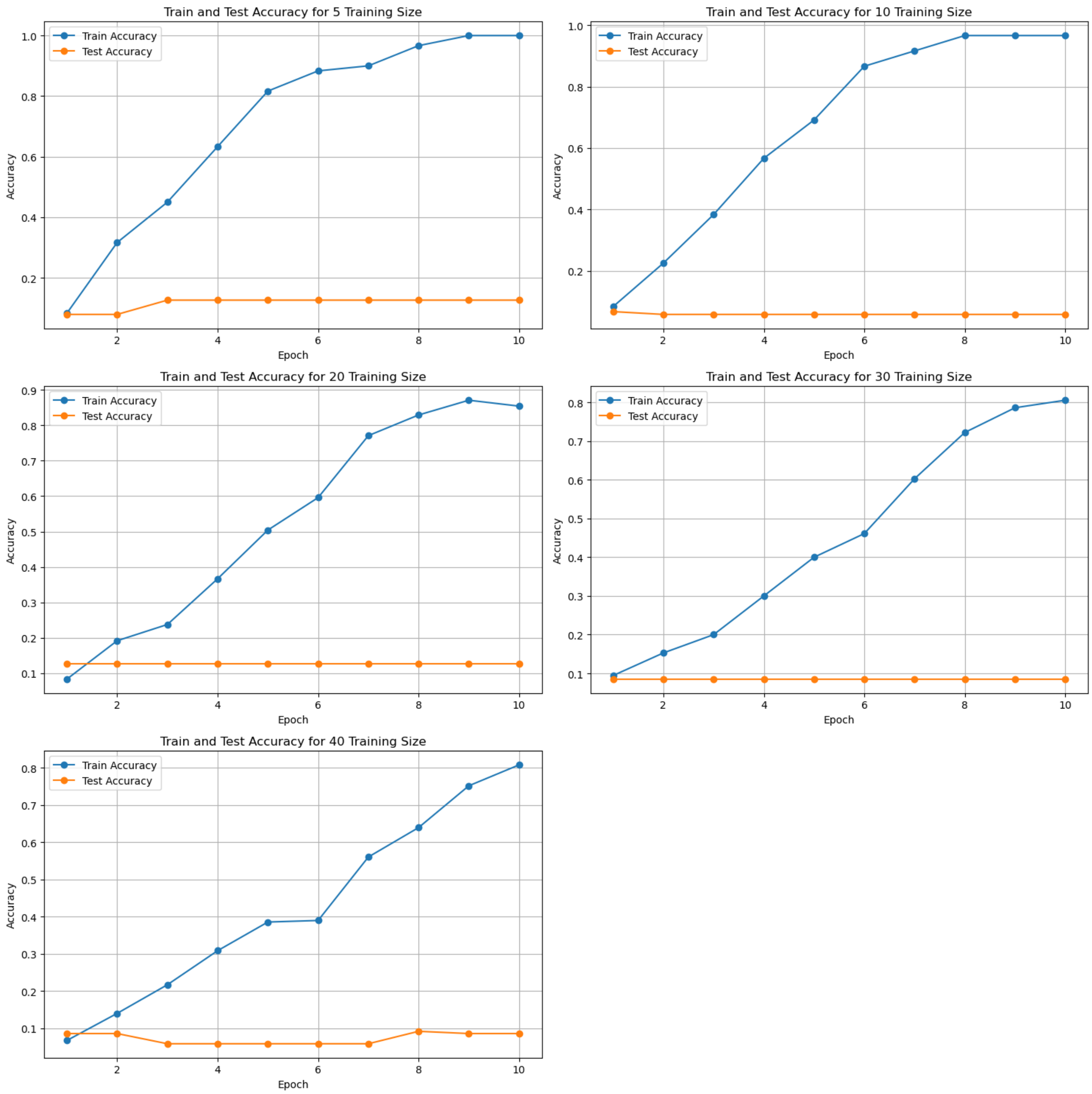
```python
        # save model
        if save_base_name:
            if not os.path.exists(save_base_name):
                os.makedirs(save_base_name)
            torch.save(model.state_dict(), f'{save_base_name}{i}.pth')

        train_size_train_accuracy.append(epoch_train_accuracy)
        train_size_test_accuracy.append(epoch_test_accuracy)

    return train_size_train_accuracy, train_size_test_accuracy
```

Create a function to display the train and test accuracies of the models

In [10]:
```python
def display_graphs(train_size_train_accuracy, train_size_test_accuracy, train_size = [5, 10, 20, 30, 40], epochs = 10):
    """
    displays graphical summaries of the train and test accuracy of train_size_train_accuracy and train_size_test_accuracy
    where train_size_test_accuracy and train_size_train_accuracy are of the same shape (5 x 10)
    train_size is the train dataset sizes used to train/test the model
    epochs is the number of epochs used
    """
    plt.figure(figsize=(15, 15))

    for idx, t_size in enumerate(train_size):
        subplot = 320 + idx + 1

        plt.subplot(subplot)
        plt.plot(list(range(1, epochs + 1)), train_size_train_accuracy[idx], label='Train Accuracy', marker='o')

        plt.plot(list(range(1, epochs + 1)), train_size_test_accuracy[idx], label='Test Accuracy', marker='o')

        plt.xlabel('Epoch')
        plt.ylabel('Accuracy')
        plt.title(f'Train and Test Accuracy for {t_size} Training Size')
        plt.legend()
        plt.grid(True)

    plt.tight_layout()
    plt.show()
```

Now we test transfer learning applicatios on the small and large model, and compare them to each other and also to the model trained with random initialization instead of transfer learning.

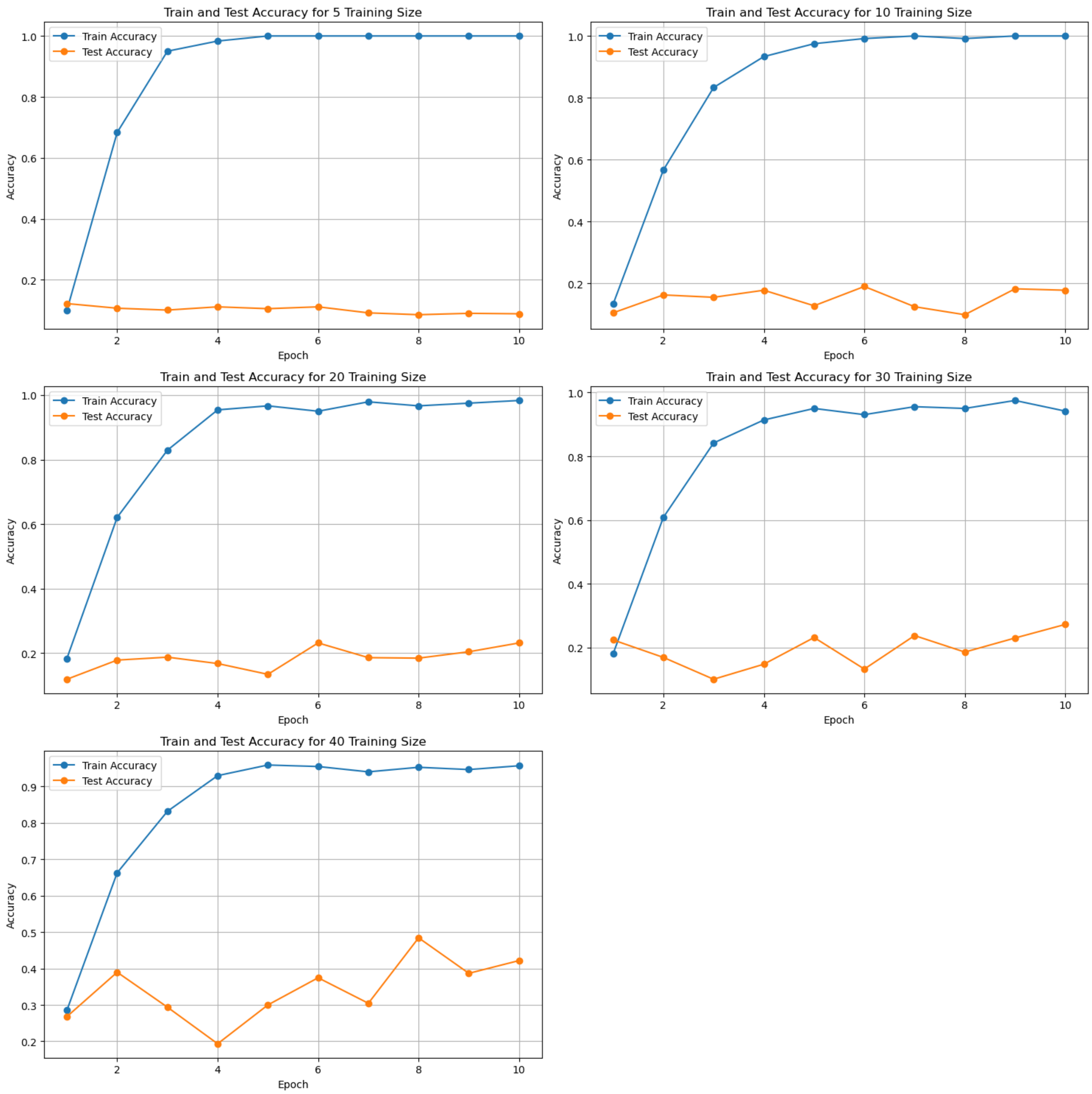
In [12]:
```python
# test small model performance with random initializatoins
train_size_train_accuracy_blank, train_size_test_accuracy_blank = train_test_sequence(model_name = "small_blank", save_base_name = "models/Bollywood/b
display_graphs(train_size_train_accuracy_blank, train_size_test_accuracy_blank)

np_array = np.array(train_size_test_accuracy_blank)
max_index = np.unravel_index(np.argmax(np_array), np_array.shape)
mean = np.mean(np_array)
print("max:", max(max(row) for row in train_size_test_accuracy_blank), "occuring at index:", max_index)
print("average:", mean)
```




```
max: 0.1263318112633181 occuring at index: (0, 2)
average: 0.09187214611872144
```

In [13]:

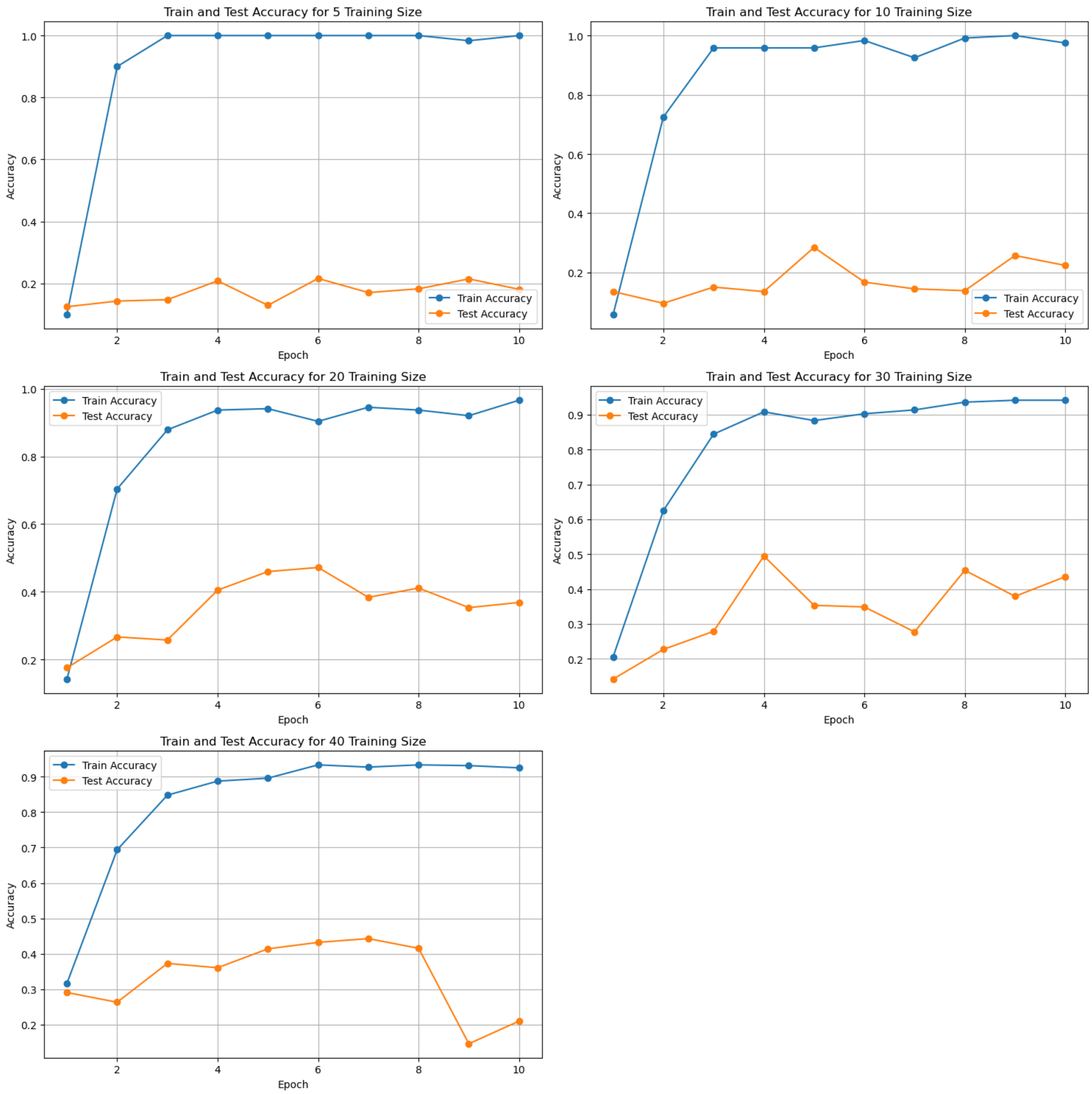
```python
# test small model performance of transfer learning
train_size_train_accuracy_small, train_size_test_accuracy_small = train_test_sequence(model_name = "small", save_base_name = "models/Bollywood/small_m
display_graphs(train_size_train_accuracy_small, train_size_test_accuracy_small)

np_array = np.array(train_size_test_accuracy_small)
max_index = np.unravel_index(np.argmax(np_array), np_array.shape)
mean = np.mean(np_array)
print("max:", max(max(row) for row in train_size_test_accuracy_small), "occuring at index:", max_index)
print("average:", mean)
```




```
max: 0.4840182648401826 occuring at index: (4, 7)
average: 0.19363774733637748
```

In [14]:

```python
# test large model performance of transfer learning
train_size_train_accuracy_large, train_size_test_accuracy_large = train_test_sequence(model_name = "large", save_base_name = "models/Bollywood/large_m
display_graphs(train_size_train_accuracy_large, train_size_test_accuracy_large)

np_array = np.array(train_size_test_accuracy_large)
max_index = np.unravel_index(np.argmax(np_array), np_array.shape)
mean = np.mean(np_array)
print("max:", max(max(row) for row in train_size_test_accuracy_large), "occuring at index:", max_index)
print("average:", mean)
```




```
max: 0.4946727549467275 occuring at index: (3, 3)
average: 0.27479452054794523
```

See the READ_ME_FIRST document for a breakdown of results.

```
In [ ]:
```