Adaption to new classes

project 7 by Guanzhao Wang, Haochen Wu, Yukai Wang

Abstact outline:

In this project, the goal is to find a way for supervised classification neural network to generalize on unlabeled data with unseen features. In particular, among all possible labels for a specific dataset, some of them do not have any labeled entries. This notebook provides a comparison of several methods we attempt to train the network utilizing the unlabeled data so that it can generalize to classifying more categories.

We use Fashion-MNIST dataset and manually remove the labels for certain categories to simulate the scenario. Then we utilize those unlabeled data by applying several unsupervised clustering methods. Such methods are implemented in `mylibs` and compared at the end of the notebook. The classification model will be trained with both labeled data and unlabeled data labeled by the clustering methods. We conclude that some of the approaches can significantly improve the models which are comparable to fully supervised models.

Teammates:

- Guanzhao Wang: Implement clustering algorithms and Variational Auto Encoder
- Haochen Wu: Implement whole flow training, dataloaders, and report plots
- Yukai Wang: Implement training models, transforms and documentation

Library

Several library functions are implemented. They are imported in the front of the next code block. The externel libraries are imported in the behind of the next block.

Our own code library is implemented in the *mylib* subdirectory. All the files in the library are listed below in an order used to train models:

- **dataloader**:

> Include customFashionMNIST class for loading custom Fashion-MNIST dataset. Dataloader functions:
> `getTrainValidateLoaders(include_labels=range(10), transform=None, batch_size=64, split = 0.9, num_workers=1, mode=7, USE_GPU=False)`
> `getTestLoaders(include_labels=range(10), transform=None, batch_size=64, num_workers=1, USE_GPU=False)`

- **transform**:

> Define transform used for dataloader.

- **model**:

> Include several Neural Networks. A normal CNN: `Net()`,
> A ResNet model: `CustomFashionResNet()`,
> clustering model: `Autoencoder()`

- **loss**:

> Include loss functions.
> crossEntropyLoss for training: `loss_function()`;
> loss for autoencoder: `autoencoder_loss()`

- **train**:

> Train functions for models and autoencoder:
> `train(train_val_loaders, net, loss_function, optimizer, USE_GPU, checkpoint_path)` and
> `autoencoder_train(train_loader, net, loss_function, optimizer, USE_GPU)`

- **eval**:

> Include validation function: `validate(val_loader, net, loss, USE_GPU)`

- **clustering**:

> Unsupervised method for labeling data. `label_data(unlabeled_data, labels, mode=0, USE_GPU=False)`

- **report**:

> Three report functions.
> Training report: `report_epoch_summary(eval_metrics)`;
> Summary report: `report_summary(mode_metrics, mode_description)`;
> Test report: `report_test_summary(mode_metrics, mode_description)`

Externel libraries included:

- **numpy**:
- **matplotlib**:

> `matplotlib.pyplt`

- **torch**:
- **torchvision**:
- **sklearn**:

> in clustering, for Kmean and GaussianMixture, confusion_matrix

In [1]:
```
from mylibs.dataloader import getTrainValidateLoaders, getTestLoaders, CustomFashionMNIST
from mylibs.train import train
from mylibs.eval import validate
from mylibs.model import Net, CustomFashionResNet
from mylibs.loss import loss_function
from mylibs.report import report_epoch_summary, report_summary, report_test_summary
from mylibs.transform import transform_t, transform_n, transform_aug
from mylibs.clustering import label_data

import numpy as np
import torch
import torchvision
from torchvision import transforms
import os
```

In [2]:
```
USE_GPU = True
BATCH_SIZE = 64
EPOCH = 30
NUM_WORKERS = 2
K = 7
```

In [3]:
```
device = torch.device("cuda" if USE_GPU else "cpu")

transform = transform_t

mode_description = {0: "clustering: kmeans",
                    1: "clustering: kmeans with PCA",
                    2: "clustering: kmeans with Auto Encoder",
                    3: "clustering: Gaussian Mixture",
                    4: "clustering: Gaussian Mixture with PCA",
                    5: "clustering: Gaussian Mixture with Auto Encoder",
                    6: "use only labeled data",
                    7: "use full FasionMNIST data",
                   }

mode_description_short = {0: "Kmeans",
                          1: "Kmeans with PCA",
                          2: "Kmeans with Auto Encoder",
                          3: "Gaussian Mixture",
                          4: "Gaussian Mixture with PCA",
                          5: "Gaussian Mixture with Auto Encoder",
                          6: "Labeled data only",
                          7: "Full FasionMNIST",
                         }

mode_metrics = {}
mode_test_metrics = {}
```

In [26]:
```
def whole_flow(mode, useResnet):
    global mode_metrics
    print(f"Getting train and validate dataloaders for mode {mode}: {mode_description[mode]}")
    train_val_loaders = getTrainValidateLoaders(include_labels=range(K), transform=transform, batch_size=BATCH_
    if useResnet:
        model = CustomFashionResNet(color_scale = 1, num_classes = 10).to(device)
    else:
        model = Net().to(device)
    optimizer = torch.optim.Adadelta(model.parameters(), lr=0.01)
    eval_metrics = []

    model_name = "ResNet" if useResnet else "Net"
    checkpoint_path = f"./checkpoint/mode_{mode}/{model_name}"
    os.makedirs(checkpoint_path, exist_ok=True)

    print(f"Start Training... {model_name}")
    # scheduler = StepLR(optimizer, step_size=1, gamma=args.gamma)
    for epoch in range(1, EPOCH+1):
```

```
        eval_metric = train(train_val_loaders, model, loss_function, optimizer, USE_GPU, f"{checkpoint_path}/ep
        eval_metrics.append(eval_metric)
        print(f"Epoch: {epoch}")
        print(f"\tTrain      - Loss: {eval_metric['train']['loss']:.4f} Accuracy: {eval_metric['train']['acc']:
        print(f"\tValidation - Loss: {eval_metric['val']['loss']:.4f} Accuracy: {eval_metric['val']['acc']:.4f}

    report_epoch_summary(eval_metrics)

    all_val_f1 = [x['val']['f1'] for x in eval_metrics]
    best_epoch = all_val_f1.index(max(all_val_f1)) + 1
    print(f"Loading model at epoch {best_epoch} for best validation f1")
    checkpoint = torch.load(f"{checkpoint_path}/epoch_{epoch}.pt")
    model.load_state_dict(checkpoint['model_state_dict'])
    optimizer.load_state_dict(checkpoint['optimizer_state_dict'])

    print("Preparing test loaders")
    labeled_test_loader, unlabeled_test_loader, test_loader = getTestLoaders(include_labels=range(K), transform

    model.eval()
    eval_metric = validate(labeled_test_loader, model, loss_function, USE_GPU)
    print(f"Result on labelled test set  : Loss: {eval_metric['loss']:.4f} Accuracy: {eval_metric['acc']:.4f} F

    eval_metric = validate(unlabeled_test_loader, model, loss_function, USE_GPU)
    print(f"Result on unlabelled test set: Loss: {eval_metric['loss']:.4f} Accuracy: {eval_metric['acc']:.4f} F

    eval_metric = validate(test_loader, model, loss_function, USE_GPU)
    print(f"Result on full test set      : Loss: {eval_metric['loss']:.4f} Accuracy: {eval_metric['acc']:.4f} F

    mode_metrics[mode] = eval_metrics
    mode_test_metrics[mode] = eval_metric

def clustering_acc():
    include_labels = range(K)
    exclude_labels = np.arange(10)[~np.isin(np.arange(10), include_labels)]

    test_set = CustomFashionMNIST(train=False, include_labels=range(10), transform=transform, mode=7,USE_GPU=US
    labeled_test_set = [(img,label) for img, label in test_set if label in include_labels]
    unlabeled_test_set = [(img,label) for img, label in test_set if label not in include_labels]
    data = [(img,label) for img, label in test_set]

    print("- Unsupervised Clusteriing on the full test set:")
    full = label_data(data, labels=np.arange(10), all_data=test_set, mode=0, USE_GPU=USE_GPU)
    print("- Unsupervised Clusteriing on the labelled test set:")
    labelled = label_data(labeled_test_set, labels=np.arange(K), all_data=test_set, mode=0, USE_GPU=USE_GPU)
```

```
    print("- Unsupervised Clusteriing on the unlabelled test set:")
    unlabelled = label_data(unlabeled_test_set, labels=exclude_labels, all_data=test_set, mode=0, USE_GPU=USE_G
```

Part 1: Basic CNN model training

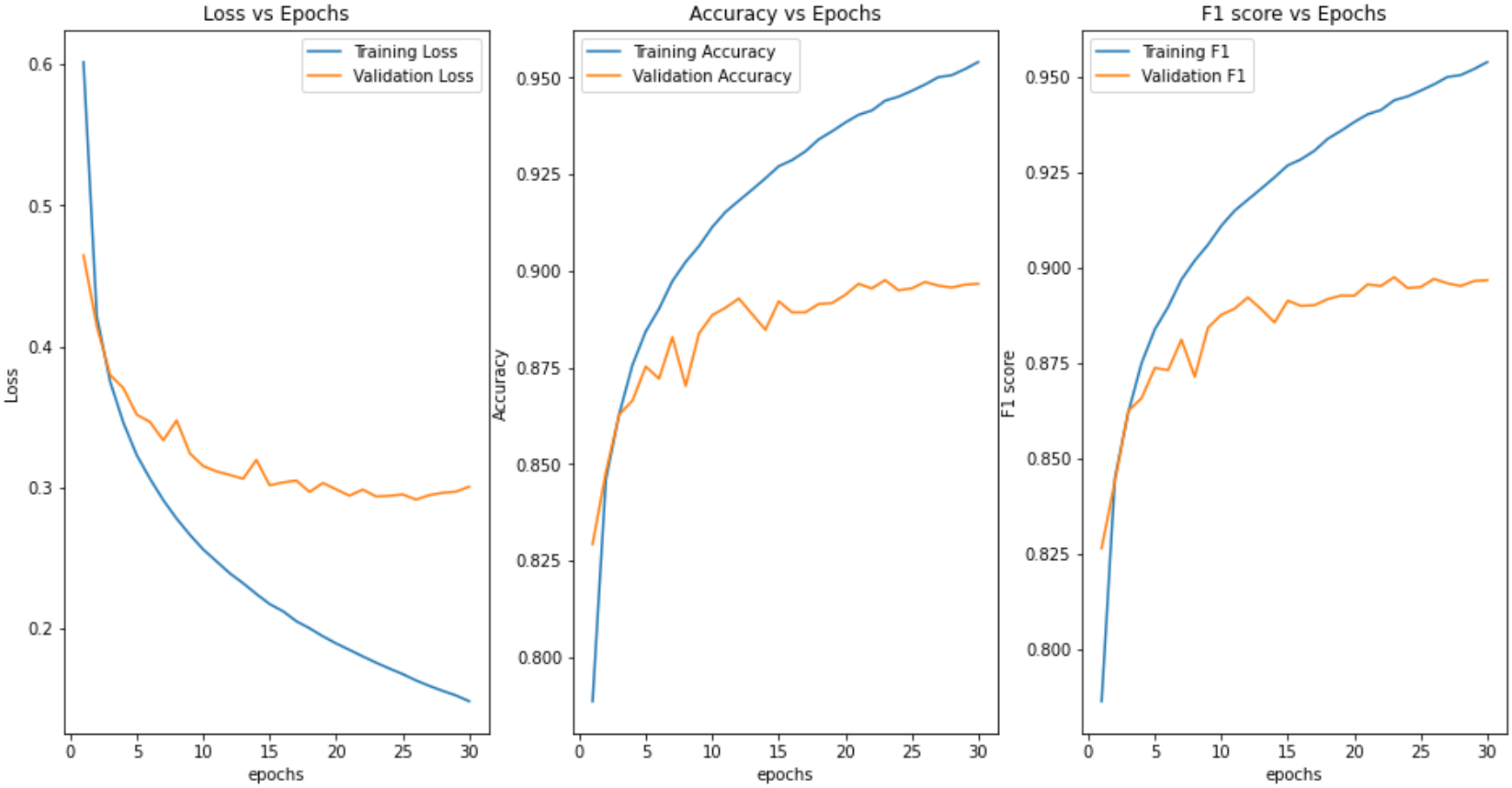
Baseline #1, use only first K-class labelled data to train

In [5]:
```
whole_flow(6, False)
```

```
Getting train and validate dataloaders for mode 6: use only labeled data
Start Training... Net
Epoch: 1
        Train      - Loss: 0.6015 Accuracy: 0.7887 F1_score: 0.7864
        Validation - Loss: 0.4646 Accuracy: 0.8293 F1_score: 0.8265
Epoch: 2
        Train      - Loss: 0.4213 Accuracy: 0.8460 F1_score: 0.8449
        Validation - Loss: 0.4144 Accuracy: 0.8479 F1_score: 0.8441
Epoch: 3
        Train      - Loss: 0.3751 Accuracy: 0.8629 F1_score: 0.8620
        Validation - Loss: 0.3797 Accuracy: 0.8629 F1_score: 0.8625
Epoch: 4
        Train      - Loss: 0.3458 Accuracy: 0.8757 F1_score: 0.8750
        Validation - Loss: 0.3704 Accuracy: 0.8664 F1_score: 0.8658
Epoch: 5
        Train      - Loss: 0.3230 Accuracy: 0.8844 F1_score: 0.8839
        Validation - Loss: 0.3516 Accuracy: 0.8752 F1_score: 0.8738
Epoch: 6
        Train      - Loss: 0.3062 Accuracy: 0.8902 F1_score: 0.8897
        Validation - Loss: 0.3464 Accuracy: 0.8721 F1_score: 0.8732
Epoch: 7
        Train      - Loss: 0.2910 Accuracy: 0.8974 F1_score: 0.8969
        Validation - Loss: 0.3334 Accuracy: 0.8829 F1_score: 0.8812
Epoch: 8
        Train      - Loss: 0.2779 Accuracy: 0.9023 F1_score: 0.9019
        Validation - Loss: 0.3475 Accuracy: 0.8702 F1_score: 0.8715
Epoch: 9
        Train      - Loss: 0.2664 Accuracy: 0.9064 F1_score: 0.9061
        Validation - Loss: 0.3241 Accuracy: 0.8838 F1_score: 0.8843
Epoch: 10
        Train      - Loss: 0.2561 Accuracy: 0.9113 F1_score: 0.9110
        Validation - Loss: 0.3151 Accuracy: 0.8886 F1_score: 0.8877
Epoch: 11
        Train      - Loss: 0.2476 Accuracy: 0.9153 F1_score: 0.9150
        Validation - Loss: 0.3114 Accuracy: 0.8905 F1_score: 0.8893
Epoch: 12
        Train      - Loss: 0.2392 Accuracy: 0.9182 F1_score: 0.9179
```

```
        Validation - Loss: 0.3089 Accuracy: 0.8929 F1_score: 0.8923
Epoch: 13
        Train      - Loss: 0.2322 Accuracy: 0.9210 F1_score: 0.9207
        Validation - Loss: 0.3061 Accuracy: 0.8888 F1_score: 0.8892
Epoch: 14
        Train      - Loss: 0.2245 Accuracy: 0.9240 F1_score: 0.9237
        Validation - Loss: 0.3196 Accuracy: 0.8848 F1_score: 0.8857
Epoch: 15
        Train      - Loss: 0.2173 Accuracy: 0.9271 F1_score: 0.9269
        Validation - Loss: 0.3015 Accuracy: 0.8921 F1_score: 0.8914
Epoch: 16
        Train      - Loss: 0.2123 Accuracy: 0.9287 F1_score: 0.9284
        Validation - Loss: 0.3036 Accuracy: 0.8893 F1_score: 0.8900
Epoch: 17
        Train      - Loss: 0.2052 Accuracy: 0.9309 F1_score: 0.9307
        Validation - Loss: 0.3049 Accuracy: 0.8893 F1_score: 0.8902
Epoch: 18
        Train      - Loss: 0.2002 Accuracy: 0.9340 F1_score: 0.9338
        Validation - Loss: 0.2968 Accuracy: 0.8914 F1_score: 0.8918
Epoch: 19
        Train      - Loss: 0.1945 Accuracy: 0.9361 F1_score: 0.9359
        Validation - Loss: 0.3032 Accuracy: 0.8917 F1_score: 0.8927
Epoch: 20
        Train      - Loss: 0.1895 Accuracy: 0.9384 F1_score: 0.9382
        Validation - Loss: 0.2986 Accuracy: 0.8938 F1_score: 0.8927
Epoch: 21
        Train      - Loss: 0.1849 Accuracy: 0.9404 F1_score: 0.9402
        Validation - Loss: 0.2942 Accuracy: 0.8967 F1_score: 0.8957
Epoch: 22
        Train      - Loss: 0.1803 Accuracy: 0.9415 F1_score: 0.9413
        Validation - Loss: 0.2985 Accuracy: 0.8955 F1_score: 0.8952
Epoch: 23
        Train      - Loss: 0.1758 Accuracy: 0.9440 F1_score: 0.9439
        Validation - Loss: 0.2936 Accuracy: 0.8976 F1_score: 0.8976
Epoch: 24
        Train      - Loss: 0.1717 Accuracy: 0.9451 F1_score: 0.9449
        Validation - Loss: 0.2941 Accuracy: 0.8950 F1_score: 0.8947
Epoch: 25
        Train      - Loss: 0.1677 Accuracy: 0.9466 F1_score: 0.9464
        Validation - Loss: 0.2951 Accuracy: 0.8955 F1_score: 0.8950
Epoch: 26
        Train      - Loss: 0.1632 Accuracy: 0.9482 F1_score: 0.9481
        Validation - Loss: 0.2914 Accuracy: 0.8971 F1_score: 0.8971
Epoch: 27
        Train      - Loss: 0.1594 Accuracy: 0.9501 F1_score: 0.9500
        Validation - Loss: 0.2946 Accuracy: 0.8962 F1_score: 0.8959
Epoch: 28
        Train      - Loss: 0.1559 Accuracy: 0.9506 F1_score: 0.9505
        Validation - Loss: 0.2962 Accuracy: 0.8957 F1_score: 0.8952
```

```
Epoch: 29
        Train      - Loss: 0.1526 Accuracy: 0.9522 F1_score: 0.9521
        Validation - Loss: 0.2970 Accuracy: 0.8964 F1_score: 0.8965
Epoch: 30
        Train      - Loss: 0.1485 Accuracy: 0.9540 F1_score: 0.9539
        Validation - Loss: 0.3005 Accuracy: 0.8967 F1_score: 0.8968
<Figure size 432x288 with 0 Axes>
```



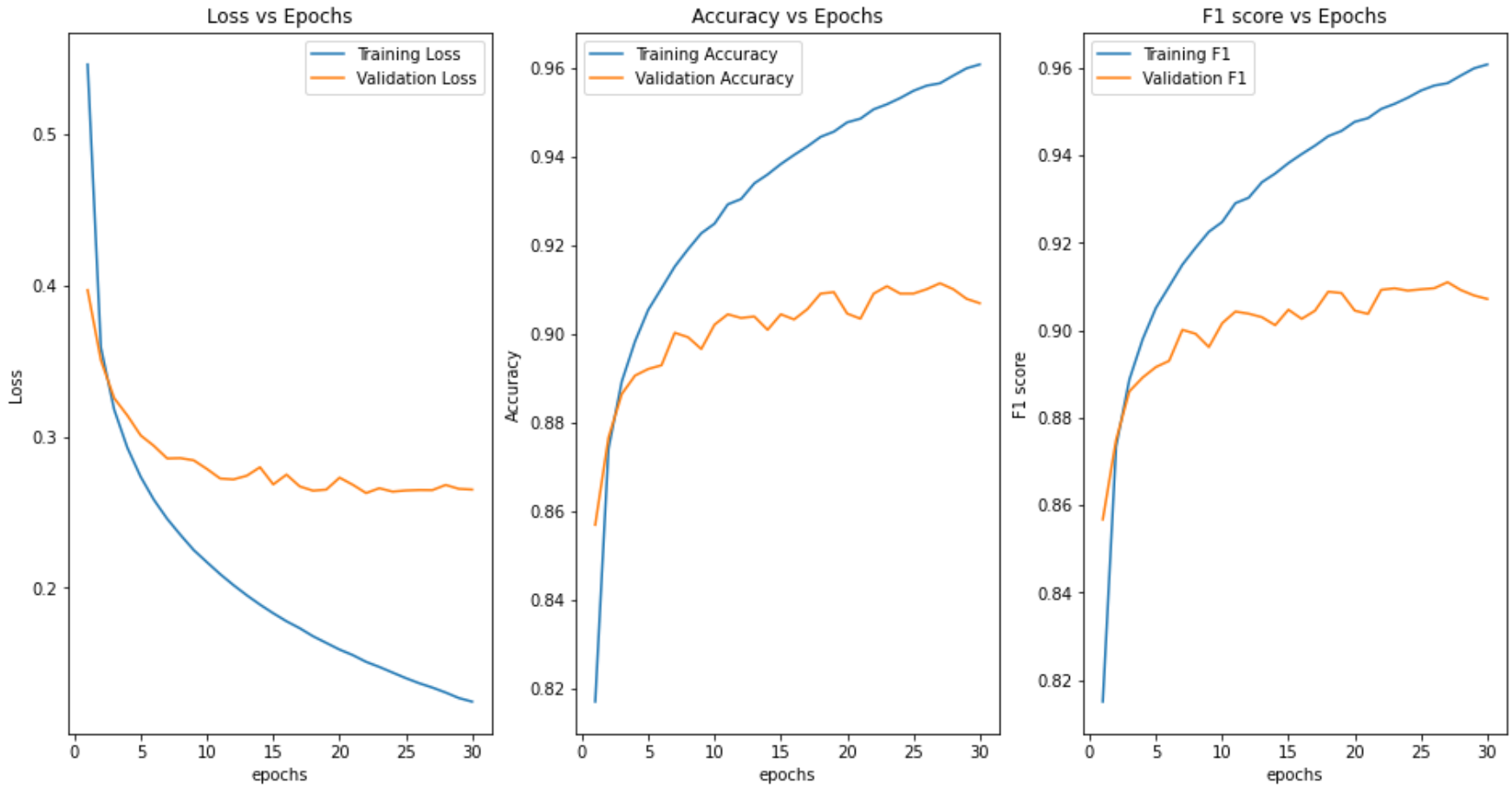

```
Loading model at epoch 23 for best validation f1
Preparing test loaders
Result on labelled test set   : Loss: 0.3275 Accuracy: 0.8856 F1_score: 0.8856
Result on unlabelled test set: Loss: 10.4008 Accuracy: 0.0000 F1_score: 0.0000
Result on full test set       : Loss: 3.3507 Accuracy: 0.6199 F1_score: 0.5362
```

Baseline #2, use full Fashion-MNIST dataset to train

In [6]:
```
whole_flow(7, False)
```

```
Getting train and validate dataloaders for mode 7: use full FasionMNIST data
Start Training... Net
Epoch: 1
        Train      - Loss: 0.5461 Accuracy: 0.8171 F1_score: 0.8151
        Validation - Loss: 0.3969 Accuracy: 0.8570 F1_score: 0.8567
Epoch: 2
        Train      - Loss: 0.3590 Accuracy: 0.8741 F1_score: 0.8734
        Validation - Loss: 0.3506 Accuracy: 0.8767 F1_score: 0.8749
Epoch: 3
        Train      - Loss: 0.3179 Accuracy: 0.8893 F1_score: 0.8887
        Validation - Loss: 0.3255 Accuracy: 0.8865 F1_score: 0.8860
Epoch: 4
        Train      - Loss: 0.2926 Accuracy: 0.8984 F1_score: 0.8979
        Validation - Loss: 0.3139 Accuracy: 0.8907 F1_score: 0.8892
Epoch: 5
        Train      - Loss: 0.2734 Accuracy: 0.9055 F1_score: 0.9051
        Validation - Loss: 0.3008 Accuracy: 0.8922 F1_score: 0.8916
Epoch: 6
        Train      - Loss: 0.2582 Accuracy: 0.9104 F1_score: 0.9100
        Validation - Loss: 0.2939 Accuracy: 0.8930 F1_score: 0.8930
Epoch: 7
        Train      - Loss: 0.2456 Accuracy: 0.9153 F1_score: 0.9150
        Validation - Loss: 0.2856 Accuracy: 0.9003 F1_score: 0.9001
Epoch: 8
        Train      - Loss: 0.2350 Accuracy: 0.9192 F1_score: 0.9189
        Validation - Loss: 0.2857 Accuracy: 0.8993 F1_score: 0.8992
Epoch: 9
        Train      - Loss: 0.2250 Accuracy: 0.9228 F1_score: 0.9225
        Validation - Loss: 0.2843 Accuracy: 0.8967 F1_score: 0.8962
Epoch: 10
        Train      - Loss: 0.2169 Accuracy: 0.9249 F1_score: 0.9248
        Validation - Loss: 0.2786 Accuracy: 0.9022 F1_score: 0.9016
Epoch: 11
        Train      - Loss: 0.2090 Accuracy: 0.9293 F1_score: 0.9291
        Validation - Loss: 0.2723 Accuracy: 0.9045 F1_score: 0.9043
Epoch: 12
        Train      - Loss: 0.2018 Accuracy: 0.9305 F1_score: 0.9303
        Validation - Loss: 0.2717 Accuracy: 0.9037 F1_score: 0.9038
Epoch: 13
        Train      - Loss: 0.1951 Accuracy: 0.9340 F1_score: 0.9339
        Validation - Loss: 0.2741 Accuracy: 0.9040 F1_score: 0.9030
Epoch: 14
        Train      - Loss: 0.1889 Accuracy: 0.9360 F1_score: 0.9358
        Validation - Loss: 0.2797 Accuracy: 0.9010 F1_score: 0.9012
Epoch: 15
        Train      - Loss: 0.1832 Accuracy: 0.9384 F1_score: 0.9382
        Validation - Loss: 0.2684 Accuracy: 0.9045 F1_score: 0.9047
Epoch: 16
        Train      - Loss: 0.1779 Accuracy: 0.9404 F1_score: 0.9403
```

```
        Validation - Loss: 0.2749 Accuracy: 0.9033 F1_score: 0.9026
Epoch: 17
        Train      - Loss: 0.1733 Accuracy: 0.9424 F1_score: 0.9422
        Validation - Loss: 0.2670 Accuracy: 0.9057 F1_score: 0.9045
Epoch: 18
        Train      - Loss: 0.1680 Accuracy: 0.9445 F1_score: 0.9444
        Validation - Loss: 0.2642 Accuracy: 0.9092 F1_score: 0.9088
Epoch: 19
        Train      - Loss: 0.1636 Accuracy: 0.9457 F1_score: 0.9456
        Validation - Loss: 0.2649 Accuracy: 0.9095 F1_score: 0.9085
Epoch: 20
        Train      - Loss: 0.1591 Accuracy: 0.9478 F1_score: 0.9477
        Validation - Loss: 0.2728 Accuracy: 0.9047 F1_score: 0.9045
Epoch: 21
        Train      - Loss: 0.1555 Accuracy: 0.9486 F1_score: 0.9485
        Validation - Loss: 0.2681 Accuracy: 0.9035 F1_score: 0.9037
Epoch: 22
        Train      - Loss: 0.1510 Accuracy: 0.9507 F1_score: 0.9506
        Validation - Loss: 0.2626 Accuracy: 0.9092 F1_score: 0.9093
Epoch: 23
        Train      - Loss: 0.1477 Accuracy: 0.9519 F1_score: 0.9518
        Validation - Loss: 0.2658 Accuracy: 0.9108 F1_score: 0.9096
Epoch: 24
        Train      - Loss: 0.1439 Accuracy: 0.9533 F1_score: 0.9532
        Validation - Loss: 0.2636 Accuracy: 0.9092 F1_score: 0.9091
Epoch: 25
        Train      - Loss: 0.1403 Accuracy: 0.9549 F1_score: 0.9548
        Validation - Loss: 0.2643 Accuracy: 0.9092 F1_score: 0.9094
Epoch: 26
        Train      - Loss: 0.1369 Accuracy: 0.9561 F1_score: 0.9560
        Validation - Loss: 0.2647 Accuracy: 0.9102 F1_score: 0.9096
Epoch: 27
        Train      - Loss: 0.1340 Accuracy: 0.9566 F1_score: 0.9565
        Validation - Loss: 0.2646 Accuracy: 0.9115 F1_score: 0.9110
Epoch: 28
        Train      - Loss: 0.1308 Accuracy: 0.9583 F1_score: 0.9582
        Validation - Loss: 0.2680 Accuracy: 0.9102 F1_score: 0.9092
Epoch: 29
        Train      - Loss: 0.1271 Accuracy: 0.9600 F1_score: 0.9599
        Validation - Loss: 0.2654 Accuracy: 0.9080 F1_score: 0.9079
Epoch: 30
        Train      - Loss: 0.1247 Accuracy: 0.9609 F1_score: 0.9608
        Validation - Loss: 0.2649 Accuracy: 0.9070 F1_score: 0.9072
<Figure size 432x288 with 0 Axes>
```




```
Loading model at epoch 27 for best validation f1
Preparing test loaders
Result on labelled test set  : Loss: 0.3431 Accuracy: 0.8801 F1_score: 0.8843
Result on unlabelled test set: Loss: 0.0965 Accuracy: 0.9697 F1_score: 0.9762
Result on full test set      : Loss: 0.2685 Accuracy: 0.9070 F1_score: 0.9073
```

Label new categories by clustering

- label the dataset under different clustering methods
- compare acuracy on the test set

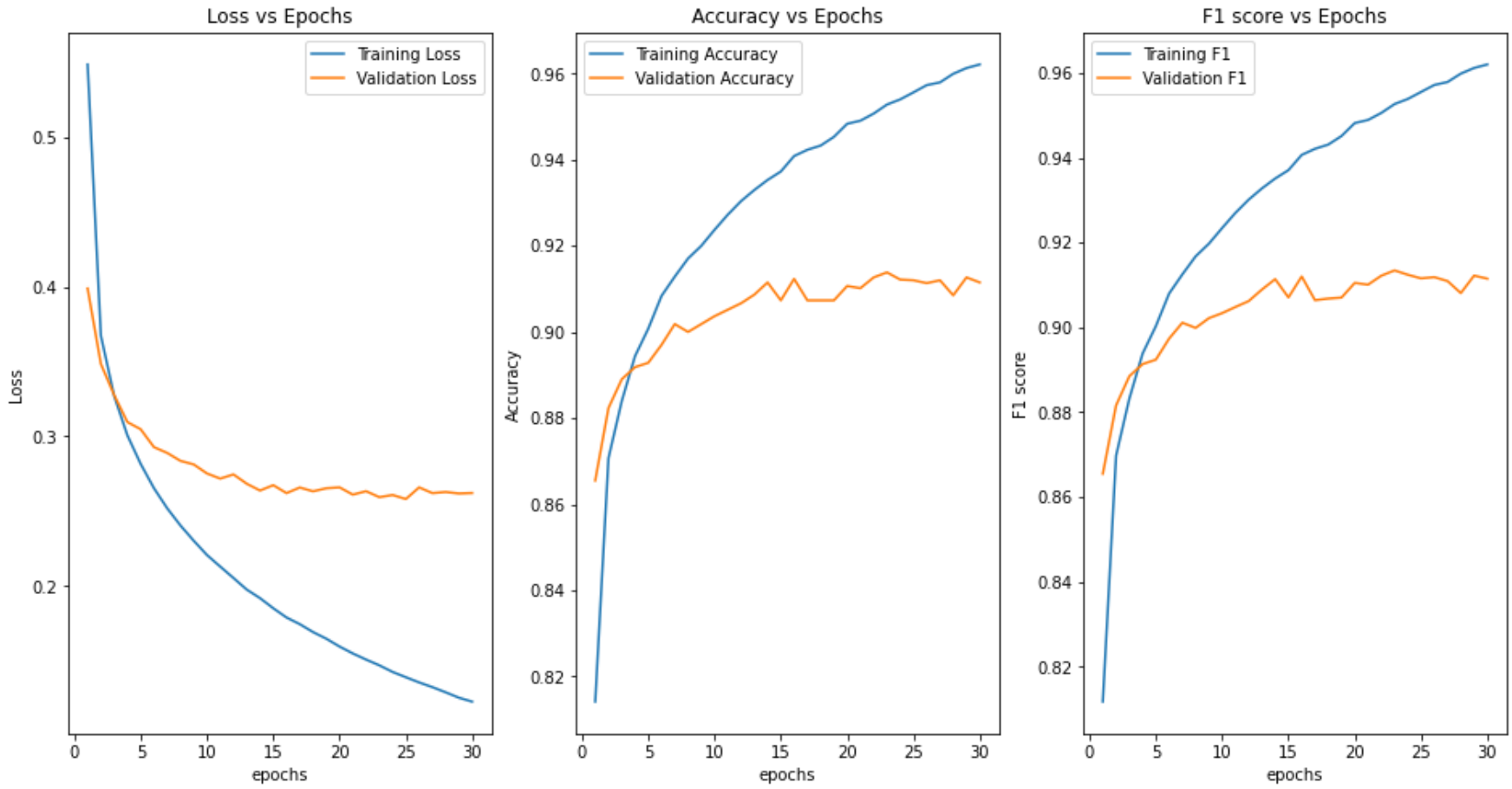
Approach #1, use KMeans to compute label for unlabelled data

```
whole_flow(0, False)
```

In [7]:

```
Getting train and validate dataloaders for mode 0: clustering: kmeans
Labeling unlabeled data...
Labeling accuracy: 0.8799444444444444
Start Training... Net
Epoch: 1
        Train      - Loss: 0.5487 Accuracy: 0.8142 F1_score: 0.8117
        Validation - Loss: 0.3990 Accuracy: 0.8655 F1_score: 0.8655
Epoch: 2
        Train      - Loss: 0.3674 Accuracy: 0.8707 F1_score: 0.8698
        Validation - Loss: 0.3485 Accuracy: 0.8823 F1_score: 0.8816
Epoch: 3
        Train      - Loss: 0.3271 Accuracy: 0.8839 F1_score: 0.8832
        Validation - Loss: 0.3277 Accuracy: 0.8890 F1_score: 0.8885
Epoch: 4
        Train      - Loss: 0.3007 Accuracy: 0.8944 F1_score: 0.8938
        Validation - Loss: 0.3096 Accuracy: 0.8918 F1_score: 0.8913
Epoch: 5
        Train      - Loss: 0.2816 Accuracy: 0.9008 F1_score: 0.9003
        Validation - Loss: 0.3048 Accuracy: 0.8928 F1_score: 0.8924
Epoch: 6
        Train      - Loss: 0.2654 Accuracy: 0.9084 F1_score: 0.9080
        Validation - Loss: 0.2928 Accuracy: 0.8970 F1_score: 0.8974
Epoch: 7
        Train      - Loss: 0.2520 Accuracy: 0.9129 F1_score: 0.9125
        Validation - Loss: 0.2890 Accuracy: 0.9018 F1_score: 0.9011
Epoch: 8
        Train      - Loss: 0.2404 Accuracy: 0.9171 F1_score: 0.9168
        Validation - Loss: 0.2836 Accuracy: 0.9000 F1_score: 0.8999
Epoch: 9
        Train      - Loss: 0.2303 Accuracy: 0.9200 F1_score: 0.9197
        Validation - Loss: 0.2813 Accuracy: 0.9018 F1_score: 0.9022
Epoch: 10
        Train      - Loss: 0.2207 Accuracy: 0.9237 F1_score: 0.9234
        Validation - Loss: 0.2752 Accuracy: 0.9037 F1_score: 0.9034
Epoch: 11
        Train      - Loss: 0.2129 Accuracy: 0.9272 F1_score: 0.9270
        Validation - Loss: 0.2718 Accuracy: 0.9052 F1_score: 0.9048
Epoch: 12
        Train      - Loss: 0.2053 Accuracy: 0.9304 F1_score: 0.9301
        Validation - Loss: 0.2746 Accuracy: 0.9067 F1_score: 0.9062
Epoch: 13
        Train      - Loss: 0.1976 Accuracy: 0.9330 F1_score: 0.9328
        Validation - Loss: 0.2684 Accuracy: 0.9087 F1_score: 0.9089
Epoch: 14
        Train      - Loss: 0.1918 Accuracy: 0.9353 F1_score: 0.9351
        Validation - Loss: 0.2638 Accuracy: 0.9115 F1_score: 0.9114
Epoch: 15
```

```
        Train      - Loss: 0.1851 Accuracy: 0.9373 F1_score: 0.9371
        Validation - Loss: 0.2674 Accuracy: 0.9073 F1_score: 0.9071
Epoch: 16
        Train      - Loss: 0.1789 Accuracy: 0.9409 F1_score: 0.9407
        Validation - Loss: 0.2621 Accuracy: 0.9123 F1_score: 0.9120
Epoch: 17
        Train      - Loss: 0.1744 Accuracy: 0.9423 F1_score: 0.9421
        Validation - Loss: 0.2658 Accuracy: 0.9073 F1_score: 0.9064
Epoch: 18
        Train      - Loss: 0.1692 Accuracy: 0.9433 F1_score: 0.9431
        Validation - Loss: 0.2633 Accuracy: 0.9073 F1_score: 0.9068
Epoch: 19
        Train      - Loss: 0.1647 Accuracy: 0.9453 F1_score: 0.9451
        Validation - Loss: 0.2653 Accuracy: 0.9073 F1_score: 0.9070
Epoch: 20
        Train      - Loss: 0.1595 Accuracy: 0.9483 F1_score: 0.9482
        Validation - Loss: 0.2659 Accuracy: 0.9107 F1_score: 0.9105
Epoch: 21
        Train      - Loss: 0.1548 Accuracy: 0.9491 F1_score: 0.9489
        Validation - Loss: 0.2611 Accuracy: 0.9102 F1_score: 0.9101
Epoch: 22
        Train      - Loss: 0.1507 Accuracy: 0.9507 F1_score: 0.9506
        Validation - Loss: 0.2634 Accuracy: 0.9127 F1_score: 0.9122
Epoch: 23
        Train      - Loss: 0.1469 Accuracy: 0.9528 F1_score: 0.9527
        Validation - Loss: 0.2594 Accuracy: 0.9138 F1_score: 0.9134
Epoch: 24
        Train      - Loss: 0.1425 Accuracy: 0.9540 F1_score: 0.9539
        Validation - Loss: 0.2608 Accuracy: 0.9122 F1_score: 0.9124
Epoch: 25
        Train      - Loss: 0.1390 Accuracy: 0.9556 F1_score: 0.9555
        Validation - Loss: 0.2581 Accuracy: 0.9120 F1_score: 0.9116
Epoch: 26
        Train      - Loss: 0.1355 Accuracy: 0.9573 F1_score: 0.9572
        Validation - Loss: 0.2659 Accuracy: 0.9113 F1_score: 0.9119
Epoch: 27
        Train      - Loss: 0.1323 Accuracy: 0.9579 F1_score: 0.9579
        Validation - Loss: 0.2621 Accuracy: 0.9120 F1_score: 0.9109
Epoch: 28
        Train      - Loss: 0.1289 Accuracy: 0.9599 F1_score: 0.9599
        Validation - Loss: 0.2628 Accuracy: 0.9085 F1_score: 0.9081
Epoch: 29
        Train      - Loss: 0.1252 Accuracy: 0.9613 F1_score: 0.9612
        Validation - Loss: 0.2618 Accuracy: 0.9127 F1_score: 0.9122
Epoch: 30
        Train      - Loss: 0.1226 Accuracy: 0.9621 F1_score: 0.9620
        Validation - Loss: 0.2621 Accuracy: 0.9115 F1_score: 0.9115
<Figure size 432x288 with 0 Axes>
```




```
Loading model at epoch 23 for best validation f1
Preparing test loaders
Result on labelled test set  : Loss: 0.3499 Accuracy: 0.8784 F1_score: 0.8829
Result on unlabelled test set: Loss: 0.7429 Accuracy: 0.8683 F1_score: 0.8772
Result on full test set      : Loss: 0.4662 Accuracy: 0.8754 F1_score: 0.8763
```

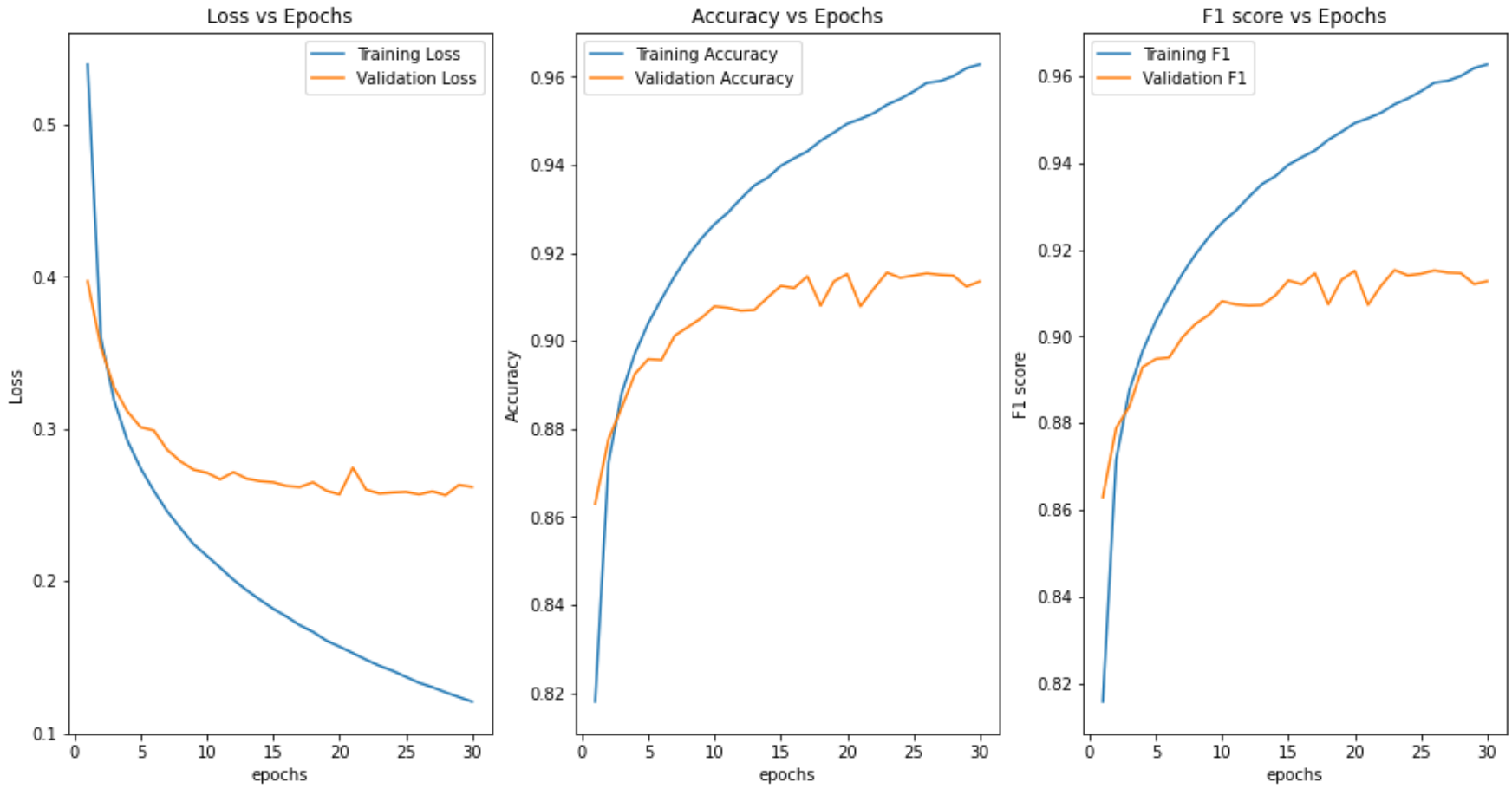
Approach #2, use KMeans with PCA to compute label for unlabelled data

In [8]:
```
whole_flow(1, False)
```

```
Getting train and validate dataloaders for mode 1: clustering: kmeans with PCA
Labeling unlabeled data...
Labeling accuracy: 0.881
Start Training... Net
Epoch: 1
        Train      - Loss: 0.5396 Accuracy: 0.8181 F1_score: 0.8158
        Validation - Loss: 0.3972 Accuracy: 0.8630 F1_score: 0.8629
```

```
Epoch: 2
        Train      - Loss: 0.3599 Accuracy: 0.8724 F1_score: 0.8715
        Validation - Loss: 0.3537 Accuracy: 0.8777 F1_score: 0.8788
Epoch: 3
        Train      - Loss: 0.3186 Accuracy: 0.8882 F1_score: 0.8876
        Validation - Loss: 0.3272 Accuracy: 0.8848 F1_score: 0.8838
Epoch: 4
        Train      - Loss: 0.2924 Accuracy: 0.8971 F1_score: 0.8966
        Validation - Loss: 0.3115 Accuracy: 0.8925 F1_score: 0.8929
Epoch: 5
        Train      - Loss: 0.2741 Accuracy: 0.9040 F1_score: 0.9035
        Validation - Loss: 0.3011 Accuracy: 0.8958 F1_score: 0.8948
Epoch: 6
        Train      - Loss: 0.2592 Accuracy: 0.9095 F1_score: 0.9092
        Validation - Loss: 0.2990 Accuracy: 0.8957 F1_score: 0.8951
Epoch: 7
        Train      - Loss: 0.2458 Accuracy: 0.9147 F1_score: 0.9144
        Validation - Loss: 0.2863 Accuracy: 0.9012 F1_score: 0.8997
Epoch: 8
        Train      - Loss: 0.2348 Accuracy: 0.9193 F1_score: 0.9189
        Validation - Loss: 0.2787 Accuracy: 0.9032 F1_score: 0.9029
Epoch: 9
        Train      - Loss: 0.2241 Accuracy: 0.9232 F1_score: 0.9229
        Validation - Loss: 0.2732 Accuracy: 0.9052 F1_score: 0.9050
Epoch: 10
        Train      - Loss: 0.2165 Accuracy: 0.9265 F1_score: 0.9262
        Validation - Loss: 0.2712 Accuracy: 0.9078 F1_score: 0.9081
Epoch: 11
        Train      - Loss: 0.2088 Accuracy: 0.9291 F1_score: 0.9289
        Validation - Loss: 0.2668 Accuracy: 0.9075 F1_score: 0.9073
Epoch: 12
        Train      - Loss: 0.2008 Accuracy: 0.9323 F1_score: 0.9321
        Validation - Loss: 0.2716 Accuracy: 0.9068 F1_score: 0.9071
Epoch: 13
        Train      - Loss: 0.1939 Accuracy: 0.9353 F1_score: 0.9351
        Validation - Loss: 0.2673 Accuracy: 0.9070 F1_score: 0.9072
Epoch: 14
        Train      - Loss: 0.1877 Accuracy: 0.9370 F1_score: 0.9369
        Validation - Loss: 0.2656 Accuracy: 0.9098 F1_score: 0.9094
Epoch: 15
        Train      - Loss: 0.1817 Accuracy: 0.9397 F1_score: 0.9396
        Validation - Loss: 0.2649 Accuracy: 0.9125 F1_score: 0.9129
Epoch: 16
        Train      - Loss: 0.1766 Accuracy: 0.9414 F1_score: 0.9413
        Validation - Loss: 0.2625 Accuracy: 0.9120 F1_score: 0.9120
Epoch: 17
        Train      - Loss: 0.1709 Accuracy: 0.9430 F1_score: 0.9429
        Validation - Loss: 0.2618 Accuracy: 0.9147 F1_score: 0.9145
Epoch: 18
```

```
        Train      - Loss: 0.1665 Accuracy: 0.9454 F1_score: 0.9453
        Validation - Loss: 0.2649 Accuracy: 0.9080 F1_score: 0.9074
Epoch: 19
        Train      - Loss: 0.1608 Accuracy: 0.9473 F1_score: 0.9471
        Validation - Loss: 0.2594 Accuracy: 0.9135 F1_score: 0.9130
Epoch: 20
        Train      - Loss: 0.1568 Accuracy: 0.9493 F1_score: 0.9492
        Validation - Loss: 0.2569 Accuracy: 0.9152 F1_score: 0.9151
Epoch: 21
        Train      - Loss: 0.1526 Accuracy: 0.9504 F1_score: 0.9503
        Validation - Loss: 0.2746 Accuracy: 0.9078 F1_score: 0.9072
Epoch: 22
        Train      - Loss: 0.1483 Accuracy: 0.9517 F1_score: 0.9516
        Validation - Loss: 0.2601 Accuracy: 0.9118 F1_score: 0.9117
Epoch: 23
        Train      - Loss: 0.1443 Accuracy: 0.9536 F1_score: 0.9535
        Validation - Loss: 0.2575 Accuracy: 0.9155 F1_score: 0.9153
Epoch: 24
        Train      - Loss: 0.1410 Accuracy: 0.9549 F1_score: 0.9548
        Validation - Loss: 0.2581 Accuracy: 0.9143 F1_score: 0.9140
Epoch: 25
        Train      - Loss: 0.1370 Accuracy: 0.9566 F1_score: 0.9565
        Validation - Loss: 0.2586 Accuracy: 0.9148 F1_score: 0.9144
Epoch: 26
        Train      - Loss: 0.1330 Accuracy: 0.9586 F1_score: 0.9585
        Validation - Loss: 0.2569 Accuracy: 0.9153 F1_score: 0.9152
Epoch: 27
        Train      - Loss: 0.1302 Accuracy: 0.9590 F1_score: 0.9589
        Validation - Loss: 0.2590 Accuracy: 0.9150 F1_score: 0.9147
Epoch: 28
        Train      - Loss: 0.1268 Accuracy: 0.9601 F1_score: 0.9600
        Validation - Loss: 0.2564 Accuracy: 0.9148 F1_score: 0.9146
Epoch: 29
        Train      - Loss: 0.1237 Accuracy: 0.9619 F1_score: 0.9618
        Validation - Loss: 0.2632 Accuracy: 0.9123 F1_score: 0.9121
Epoch: 30
        Train      - Loss: 0.1207 Accuracy: 0.9628 F1_score: 0.9627
        Validation - Loss: 0.2618 Accuracy: 0.9135 F1_score: 0.9127
<Figure size 432x288 with 0 Axes>
```




```
Loading model at epoch 23 for best validation f1
Preparing test loaders
Result on labelled test set  : Loss: 0.3549 Accuracy: 0.8773 F1_score: 0.8808
Result on unlabelled test set: Loss: 0.7100 Accuracy: 0.8773 F1_score: 0.8850
Result on full test set      : Loss: 0.4614 Accuracy: 0.8773 F1_score: 0.8769
```

Approach #3, use KMeans with Auto Encoder to compute label for unlabelled data

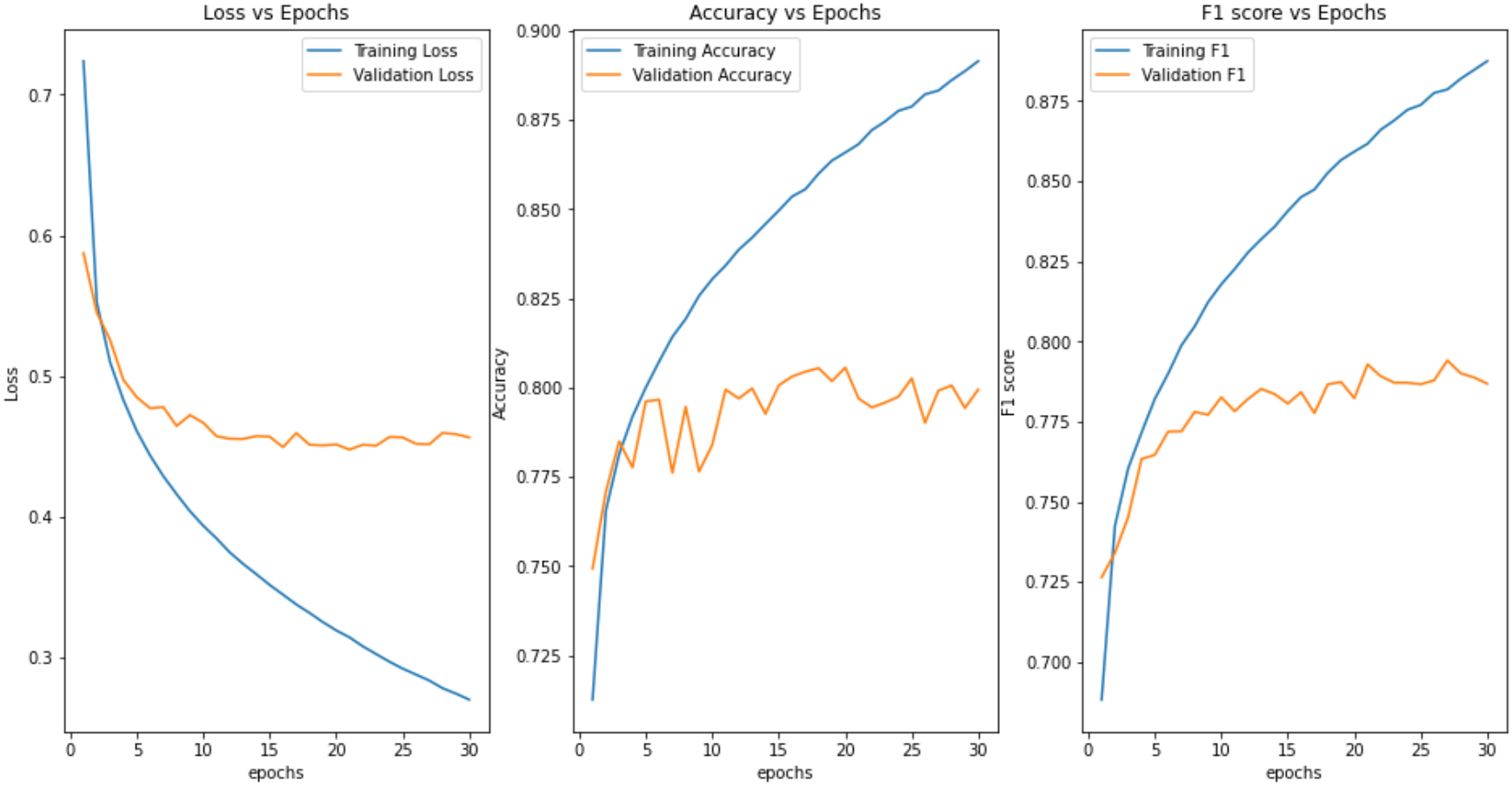
In [9]:
```
whole_flow(2, False)
```

```
Getting train and validate dataloaders for mode 2: clustering: kmeans with Auto Encoder
Labeling unlabeled data...
Training Auto Encoder...
/ssd/jason_ssd/AdaptToNewClass/mylibs/loss.py:8: UserWarning: To copy construct from a tensor, it is recommende
d to use sourceTensor.clone().detach() or sourceTensor.clone().detach().requires_grad_(True), rather than torc
```

```
h.tensor(sourceTensor).
  log2pi = torch.log(torch.tensor(2.0 * torch.as_tensor(np.pi)))
/ssd/jason_ssd/AdaptToNewClass/mylibs/loss.py:9: UserWarning: To copy construct from a tensor, it is recommende
d to use sourceTensor.clone().detach() or sourceTensor.clone().detach().requires_grad_(True), rather than torc
h.tensor(sourceTensor).
  return torch.sum(-0.5 * ((sample - mean) ** 2.0 * torch.exp(torch.tensor(-logvar)) + logvar + log2pi), dim=ra
xis)
Epoch: 1 Loss: {'loss': 354.99633662109375}
Epoch: 2 Loss: {'loss': 320.9731922526042}
Epoch: 3 Loss: {'loss': 317.61745400390623}
Epoch: 4 Loss: {'loss': 316.3578083170573}
Epoch: 5 Loss: {'loss': 315.8289954264323}
Epoch: 6 Loss: {'loss': 315.3753229329427}
Epoch: 7 Loss: {'loss': 315.15972548828125}
Epoch: 8 Loss: {'loss': 314.9283509440104}
Epoch: 9 Loss: {'loss': 314.51958159179685}
Epoch: 10 Loss: {'loss': 314.3853491210937}
Epoch: 11 Loss: {'loss': 314.3160645345052}
Epoch: 12 Loss: {'loss': 314.3321009440104}
Epoch: 13 Loss: {'loss': 314.1502100260417}
Epoch: 14 Loss: {'loss': 313.933809000651}
Epoch: 15 Loss: {'loss': 314.0860360514323}
Labeling accuracy: 0.34144444444444444
Start Training... Net
Epoch: 1
        Train      - Loss: 0.7237 Accuracy: 0.7126 F1_score: 0.6883
        Validation - Loss: 0.5872 Accuracy: 0.7493 F1_score: 0.7264
Epoch: 2
        Train      - Loss: 0.5523 Accuracy: 0.7656 F1_score: 0.7425
        Validation - Loss: 0.5450 Accuracy: 0.7708 F1_score: 0.7341
Epoch: 3
        Train      - Loss: 0.5097 Accuracy: 0.7814 F1_score: 0.7605
        Validation - Loss: 0.5256 Accuracy: 0.7850 F1_score: 0.7453
Epoch: 4
        Train      - Loss: 0.4828 Accuracy: 0.7919 F1_score: 0.7716
        Validation - Loss: 0.4971 Accuracy: 0.7777 F1_score: 0.7634
Epoch: 5
        Train      - Loss: 0.4609 Accuracy: 0.8001 F1_score: 0.7821
        Validation - Loss: 0.4849 Accuracy: 0.7962 F1_score: 0.7646
Epoch: 6
        Train      - Loss: 0.4436 Accuracy: 0.8073 F1_score: 0.7900
        Validation - Loss: 0.4771 Accuracy: 0.7967 F1_score: 0.7719
Epoch: 7
        Train      - Loss: 0.4288 Accuracy: 0.8143 F1_score: 0.7988
        Validation - Loss: 0.4779 Accuracy: 0.7762 F1_score: 0.7720
Epoch: 8
        Train      - Loss: 0.4160 Accuracy: 0.8193 F1_score: 0.8047
        Validation - Loss: 0.4644 Accuracy: 0.7947 F1_score: 0.7780
Epoch: 9
```

```
         Train      - Loss: 0.4039 Accuracy: 0.8257 F1_score: 0.8122
         Validation - Loss: 0.4722 Accuracy: 0.7765 F1_score: 0.7771
Epoch: 10
         Train      - Loss: 0.3936 Accuracy: 0.8304 F1_score: 0.8178
         Validation - Loss: 0.4668 Accuracy: 0.7840 F1_score: 0.7826
Epoch: 11
         Train      - Loss: 0.3846 Accuracy: 0.8342 F1_score: 0.8226
         Validation - Loss: 0.4572 Accuracy: 0.7995 F1_score: 0.7782
Epoch: 12
         Train      - Loss: 0.3746 Accuracy: 0.8387 F1_score: 0.8277
         Validation - Loss: 0.4554 Accuracy: 0.7970 F1_score: 0.7820
Epoch: 13
         Train      - Loss: 0.3665 Accuracy: 0.8421 F1_score: 0.8319
         Validation - Loss: 0.4552 Accuracy: 0.7998 F1_score: 0.7852
Epoch: 14
         Train      - Loss: 0.3591 Accuracy: 0.8459 F1_score: 0.8358
         Validation - Loss: 0.4573 Accuracy: 0.7927 F1_score: 0.7836
Epoch: 15
         Train      - Loss: 0.3515 Accuracy: 0.8496 F1_score: 0.8406
         Validation - Loss: 0.4569 Accuracy: 0.8007 F1_score: 0.7806
Epoch: 16
         Train      - Loss: 0.3446 Accuracy: 0.8536 F1_score: 0.8450
         Validation - Loss: 0.4494 Accuracy: 0.8032 F1_score: 0.7842
Epoch: 17
         Train      - Loss: 0.3376 Accuracy: 0.8556 F1_score: 0.8474
         Validation - Loss: 0.4594 Accuracy: 0.8045 F1_score: 0.7776
Epoch: 18
         Train      - Loss: 0.3316 Accuracy: 0.8600 F1_score: 0.8526
         Validation - Loss: 0.4511 Accuracy: 0.8055 F1_score: 0.7866
Epoch: 19
         Train      - Loss: 0.3251 Accuracy: 0.8636 F1_score: 0.8566
         Validation - Loss: 0.4506 Accuracy: 0.8018 F1_score: 0.7874
Epoch: 20
         Train      - Loss: 0.3192 Accuracy: 0.8659 F1_score: 0.8592
         Validation - Loss: 0.4513 Accuracy: 0.8057 F1_score: 0.7823
Epoch: 21
         Train      - Loss: 0.3142 Accuracy: 0.8682 F1_score: 0.8617
         Validation - Loss: 0.4477 Accuracy: 0.7970 F1_score: 0.7929
Epoch: 22
         Train      - Loss: 0.3079 Accuracy: 0.8722 F1_score: 0.8661
         Validation - Loss: 0.4511 Accuracy: 0.7945 F1_score: 0.7891
Epoch: 23
         Train      - Loss: 0.3024 Accuracy: 0.8746 F1_score: 0.8689
         Validation - Loss: 0.4504 Accuracy: 0.7958 F1_score: 0.7871
Epoch: 24
         Train      - Loss: 0.2969 Accuracy: 0.8776 F1_score: 0.8722
         Validation - Loss: 0.4566 Accuracy: 0.7975 F1_score: 0.7871
Epoch: 25
         Train      - Loss: 0.2920 Accuracy: 0.8788 F1_score: 0.8738
```

```
        Validation - Loss: 0.4564 Accuracy: 0.8027 F1_score: 0.7867
Epoch: 26
        Train      - Loss: 0.2878 Accuracy: 0.8822 F1_score: 0.8775
        Validation - Loss: 0.4518 Accuracy: 0.7902 F1_score: 0.7879
Epoch: 27
        Train      - Loss: 0.2835 Accuracy: 0.8833 F1_score: 0.8786
        Validation - Loss: 0.4514 Accuracy: 0.7992 F1_score: 0.7941
Epoch: 28
        Train      - Loss: 0.2781 Accuracy: 0.8862 F1_score: 0.8819
        Validation - Loss: 0.4596 Accuracy: 0.8007 F1_score: 0.7901
Epoch: 29
        Train      - Loss: 0.2743 Accuracy: 0.8887 F1_score: 0.8847
        Validation - Loss: 0.4586 Accuracy: 0.7943 F1_score: 0.7888
Epoch: 30
        Train      - Loss: 0.2700 Accuracy: 0.8915 F1_score: 0.8874
        Validation - Loss: 0.4563 Accuracy: 0.7995 F1_score: 0.7868
<Figure size 432x288 with 0 Axes>
```




```
Loading model at epoch 27 for best validation f1
```

```
Preparing test loaders
Result on labelled test set  : Loss: 0.3618 Accuracy: 0.8753 F1_score: 0.8790
Result on unlabelled test set: Loss: 4.3273 Accuracy: 0.3537 F1_score: 0.2550
Result on full test set      : Loss: 1.5523 Accuracy: 0.7188 F1_score: 0.6891
```

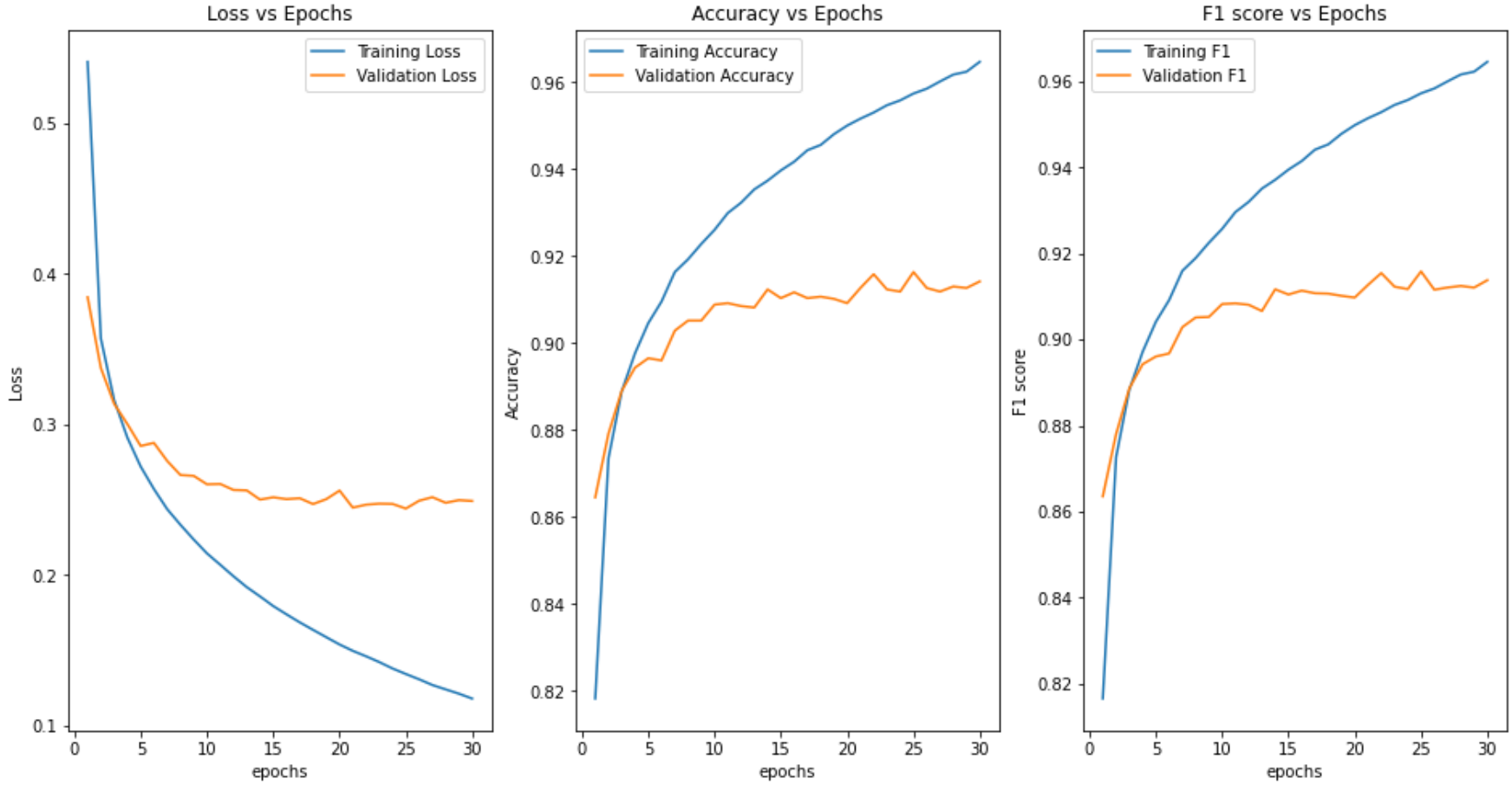
Approach #4, use Gaussian Mixture to compute label for unlabelled data

In [10]:
```
whole_flow(3, False)
```

```
Getting train and validate dataloaders for mode 3: clustering: Gaussian Mixture
Labeling unlabeled data...
Labeling accuracy: 0.8328888888888889
Start Training... Net
Epoch: 1
        Train      - Loss: 0.5408 Accuracy: 0.8182 F1_score: 0.8165
        Validation - Loss: 0.3846 Accuracy: 0.8645 F1_score: 0.8635
Epoch: 2
        Train      - Loss: 0.3572 Accuracy: 0.8735 F1_score: 0.8727
        Validation - Loss: 0.3376 Accuracy: 0.8792 F1_score: 0.8780
Epoch: 3
        Train      - Loss: 0.3163 Accuracy: 0.8889 F1_score: 0.8883
        Validation - Loss: 0.3138 Accuracy: 0.8890 F1_score: 0.8887
Epoch: 4
        Train      - Loss: 0.2913 Accuracy: 0.8976 F1_score: 0.8971
        Validation - Loss: 0.3003 Accuracy: 0.8943 F1_score: 0.8943
Epoch: 5
        Train      - Loss: 0.2722 Accuracy: 0.9046 F1_score: 0.9041
        Validation - Loss: 0.2858 Accuracy: 0.8965 F1_score: 0.8961
Epoch: 6
        Train      - Loss: 0.2572 Accuracy: 0.9095 F1_score: 0.9092
        Validation - Loss: 0.2879 Accuracy: 0.8960 F1_score: 0.8968
Epoch: 7
        Train      - Loss: 0.2439 Accuracy: 0.9163 F1_score: 0.9160
        Validation - Loss: 0.2759 Accuracy: 0.9028 F1_score: 0.9029
Epoch: 8
        Train      - Loss: 0.2335 Accuracy: 0.9193 F1_score: 0.9190
        Validation - Loss: 0.2665 Accuracy: 0.9052 F1_score: 0.9051
Epoch: 9
        Train      - Loss: 0.2237 Accuracy: 0.9228 F1_score: 0.9225
        Validation - Loss: 0.2660 Accuracy: 0.9052 F1_score: 0.9053
Epoch: 10
        Train      - Loss: 0.2145 Accuracy: 0.9260 F1_score: 0.9258
        Validation - Loss: 0.2604 Accuracy: 0.9088 F1_score: 0.9083
Epoch: 11
        Train      - Loss: 0.2070 Accuracy: 0.9299 F1_score: 0.9297
        Validation - Loss: 0.2606 Accuracy: 0.9092 F1_score: 0.9084
Epoch: 12
        Train      - Loss: 0.1993 Accuracy: 0.9323 F1_score: 0.9320
```

```
         Validation - Loss: 0.2566 Accuracy: 0.9085 F1_score: 0.9081
Epoch: 13
         Train      - Loss: 0.1921 Accuracy: 0.9354 F1_score: 0.9351
         Validation - Loss: 0.2563 Accuracy: 0.9082 F1_score: 0.9067
Epoch: 14
         Train      - Loss: 0.1859 Accuracy: 0.9374 F1_score: 0.9372
         Validation - Loss: 0.2502 Accuracy: 0.9123 F1_score: 0.9117
Epoch: 15
         Train      - Loss: 0.1795 Accuracy: 0.9397 F1_score: 0.9395
         Validation - Loss: 0.2517 Accuracy: 0.9103 F1_score: 0.9105
Epoch: 16
         Train      - Loss: 0.1740 Accuracy: 0.9417 F1_score: 0.9415
         Validation - Loss: 0.2505 Accuracy: 0.9117 F1_score: 0.9114
Epoch: 17
         Train      - Loss: 0.1687 Accuracy: 0.9444 F1_score: 0.9442
         Validation - Loss: 0.2511 Accuracy: 0.9103 F1_score: 0.9108
Epoch: 18
         Train      - Loss: 0.1637 Accuracy: 0.9456 F1_score: 0.9454
         Validation - Loss: 0.2473 Accuracy: 0.9107 F1_score: 0.9107
Epoch: 19
         Train      - Loss: 0.1589 Accuracy: 0.9480 F1_score: 0.9479
         Validation - Loss: 0.2505 Accuracy: 0.9102 F1_score: 0.9102
Epoch: 20
         Train      - Loss: 0.1540 Accuracy: 0.9500 F1_score: 0.9499
         Validation - Loss: 0.2562 Accuracy: 0.9092 F1_score: 0.9098
Epoch: 21
         Train      - Loss: 0.1498 Accuracy: 0.9516 F1_score: 0.9515
         Validation - Loss: 0.2449 Accuracy: 0.9127 F1_score: 0.9127
Epoch: 22
         Train      - Loss: 0.1462 Accuracy: 0.9530 F1_score: 0.9529
         Validation - Loss: 0.2468 Accuracy: 0.9158 F1_score: 0.9155
Epoch: 23
         Train      - Loss: 0.1423 Accuracy: 0.9547 F1_score: 0.9546
         Validation - Loss: 0.2475 Accuracy: 0.9123 F1_score: 0.9123
Epoch: 24
         Train      - Loss: 0.1381 Accuracy: 0.9558 F1_score: 0.9557
         Validation - Loss: 0.2473 Accuracy: 0.9118 F1_score: 0.9118
Epoch: 25
         Train      - Loss: 0.1344 Accuracy: 0.9574 F1_score: 0.9573
         Validation - Loss: 0.2443 Accuracy: 0.9163 F1_score: 0.9159
Epoch: 26
         Train      - Loss: 0.1309 Accuracy: 0.9585 F1_score: 0.9584
         Validation - Loss: 0.2495 Accuracy: 0.9127 F1_score: 0.9116
Epoch: 27
         Train      - Loss: 0.1271 Accuracy: 0.9602 F1_score: 0.9601
         Validation - Loss: 0.2518 Accuracy: 0.9118 F1_score: 0.9121
Epoch: 28
         Train      - Loss: 0.1242 Accuracy: 0.9617 F1_score: 0.9616
         Validation - Loss: 0.2481 Accuracy: 0.9130 F1_score: 0.9125
```

```
Epoch: 29
        Train      - Loss: 0.1214 Accuracy: 0.9624 F1_score: 0.9623
        Validation - Loss: 0.2498 Accuracy: 0.9127 F1_score: 0.9121
Epoch: 30
        Train      - Loss: 0.1181 Accuracy: 0.9647 F1_score: 0.9646
        Validation - Loss: 0.2493 Accuracy: 0.9142 F1_score: 0.9138
<Figure size 432x288 with 0 Axes>
```



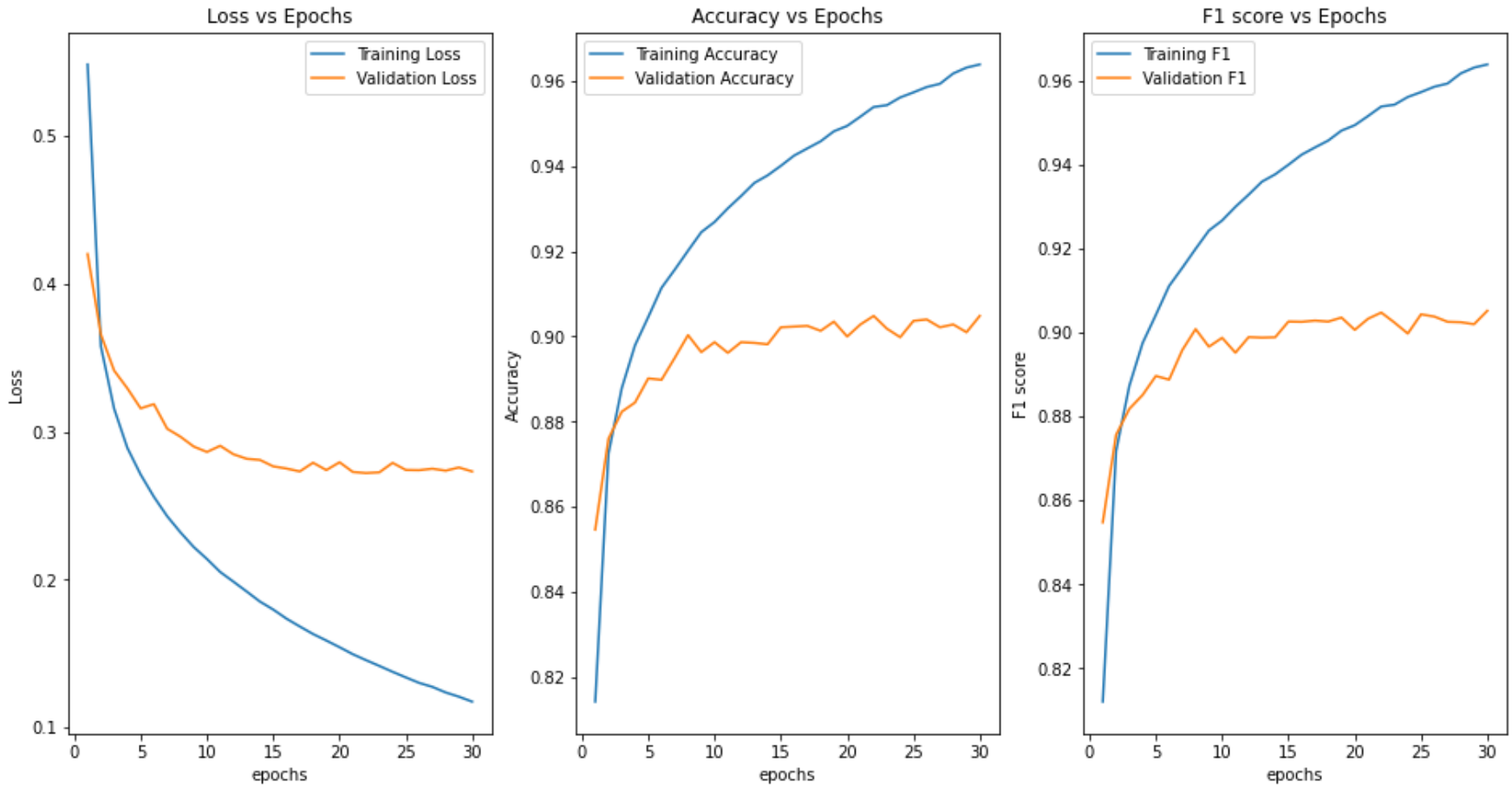

```
Loading model at epoch 25 for best validation f1
Preparing test loaders
Result on labelled test set  : Loss: 0.3434 Accuracy: 0.8786 F1_score: 0.8819
Result on unlabelled test set: Loss: 1.3362 Accuracy: 0.8250 F1_score: 0.8280
Result on full test set      : Loss: 0.6390 Accuracy: 0.8625 F1_score: 0.8614
```

Approach #5, use Gaussian Mixture with PCA to compute label for unlabelled data

In [11]:
```
whole_flow(4, False)
```

```
Getting train and validate dataloaders for mode 4: clustering: Gaussian Mixture with PCA
Labeling unlabeled data...
Labeling accuracy: 0.6936666666666667
Start Training... Net
Epoch: 1
        Train      - Loss: 0.5484 Accuracy: 0.8143 F1_score: 0.8121
        Validation - Loss: 0.4202 Accuracy: 0.8547 F1_score: 0.8548
Epoch: 2
        Train      - Loss: 0.3578 Accuracy: 0.8725 F1_score: 0.8717
        Validation - Loss: 0.3658 Accuracy: 0.8760 F1_score: 0.8756
Epoch: 3
        Train      - Loss: 0.3155 Accuracy: 0.8878 F1_score: 0.8873
        Validation - Loss: 0.3413 Accuracy: 0.8823 F1_score: 0.8818
Epoch: 4
        Train      - Loss: 0.2890 Accuracy: 0.8979 F1_score: 0.8974
        Validation - Loss: 0.3293 Accuracy: 0.8845 F1_score: 0.8851
Epoch: 5
        Train      - Loss: 0.2710 Accuracy: 0.9046 F1_score: 0.9042
        Validation - Loss: 0.3158 Accuracy: 0.8902 F1_score: 0.8896
Epoch: 6
        Train      - Loss: 0.2559 Accuracy: 0.9114 F1_score: 0.9110
        Validation - Loss: 0.3187 Accuracy: 0.8898 F1_score: 0.8887
Epoch: 7
        Train      - Loss: 0.2428 Accuracy: 0.9157 F1_score: 0.9154
        Validation - Loss: 0.3020 Accuracy: 0.8950 F1_score: 0.8958
Epoch: 8
        Train      - Loss: 0.2319 Accuracy: 0.9201 F1_score: 0.9199
        Validation - Loss: 0.2966 Accuracy: 0.9003 F1_score: 0.9008
Epoch: 9
        Train      - Loss: 0.2220 Accuracy: 0.9245 F1_score: 0.9242
        Validation - Loss: 0.2900 Accuracy: 0.8963 F1_score: 0.8966
Epoch: 10
        Train      - Loss: 0.2139 Accuracy: 0.9269 F1_score: 0.9266
        Validation - Loss: 0.2863 Accuracy: 0.8987 F1_score: 0.8987
Epoch: 11
        Train      - Loss: 0.2051 Accuracy: 0.9301 F1_score: 0.9299
        Validation - Loss: 0.2905 Accuracy: 0.8962 F1_score: 0.8952
Epoch: 12
        Train      - Loss: 0.1986 Accuracy: 0.9330 F1_score: 0.9328
        Validation - Loss: 0.2847 Accuracy: 0.8987 F1_score: 0.8989
Epoch: 13
        Train      - Loss: 0.1919 Accuracy: 0.9361 F1_score: 0.9359
        Validation - Loss: 0.2817 Accuracy: 0.8985 F1_score: 0.8987
Epoch: 14
        Train      - Loss: 0.1851 Accuracy: 0.9378 F1_score: 0.9376
        Validation - Loss: 0.2809 Accuracy: 0.8982 F1_score: 0.8988
Epoch: 15
```

```
        Train      - Loss: 0.1797 Accuracy: 0.9401 F1_score: 0.9399
        Validation - Loss: 0.2765 Accuracy: 0.9022 F1_score: 0.9026
Epoch: 16
        Train      - Loss: 0.1735 Accuracy: 0.9425 F1_score: 0.9423
        Validation - Loss: 0.2751 Accuracy: 0.9023 F1_score: 0.9025
Epoch: 17
        Train      - Loss: 0.1683 Accuracy: 0.9442 F1_score: 0.9440
        Validation - Loss: 0.2731 Accuracy: 0.9025 F1_score: 0.9028
Epoch: 18
        Train      - Loss: 0.1632 Accuracy: 0.9458 F1_score: 0.9457
        Validation - Loss: 0.2790 Accuracy: 0.9013 F1_score: 0.9026
Epoch: 19
        Train      - Loss: 0.1588 Accuracy: 0.9482 F1_score: 0.9480
        Validation - Loss: 0.2740 Accuracy: 0.9035 F1_score: 0.9035
Epoch: 20
        Train      - Loss: 0.1542 Accuracy: 0.9494 F1_score: 0.9493
        Validation - Loss: 0.2793 Accuracy: 0.9000 F1_score: 0.9006
Epoch: 21
        Train      - Loss: 0.1495 Accuracy: 0.9516 F1_score: 0.9515
        Validation - Loss: 0.2728 Accuracy: 0.9028 F1_score: 0.9032
Epoch: 22
        Train      - Loss: 0.1455 Accuracy: 0.9539 F1_score: 0.9538
        Validation - Loss: 0.2720 Accuracy: 0.9048 F1_score: 0.9047
Epoch: 23
        Train      - Loss: 0.1416 Accuracy: 0.9543 F1_score: 0.9542
        Validation - Loss: 0.2725 Accuracy: 0.9018 F1_score: 0.9023
Epoch: 24
        Train      - Loss: 0.1376 Accuracy: 0.9561 F1_score: 0.9560
        Validation - Loss: 0.2789 Accuracy: 0.8998 F1_score: 0.8997
Epoch: 25
        Train      - Loss: 0.1338 Accuracy: 0.9573 F1_score: 0.9572
        Validation - Loss: 0.2741 Accuracy: 0.9037 F1_score: 0.9043
Epoch: 26
        Train      - Loss: 0.1301 Accuracy: 0.9586 F1_score: 0.9585
        Validation - Loss: 0.2740 Accuracy: 0.9040 F1_score: 0.9038
Epoch: 27
        Train      - Loss: 0.1274 Accuracy: 0.9593 F1_score: 0.9592
        Validation - Loss: 0.2750 Accuracy: 0.9022 F1_score: 0.9025
Epoch: 28
        Train      - Loss: 0.1237 Accuracy: 0.9618 F1_score: 0.9617
        Validation - Loss: 0.2736 Accuracy: 0.9028 F1_score: 0.9024
Epoch: 29
        Train      - Loss: 0.1208 Accuracy: 0.9631 F1_score: 0.9630
        Validation - Loss: 0.2758 Accuracy: 0.9010 F1_score: 0.9019
Epoch: 30
        Train      - Loss: 0.1174 Accuracy: 0.9639 F1_score: 0.9638
        Validation - Loss: 0.2731 Accuracy: 0.9048 F1_score: 0.9051
<Figure size 432x288 with 0 Axes>
```




```
Loading model at epoch 30 for best validation f1
Preparing test loaders
Result on labelled test set  : Loss: 0.3369 Accuracy: 0.8830 F1_score: 0.8866
Result on unlabelled test set: Loss: 3.3661 Accuracy: 0.6863 F1_score: 0.6731
Result on full test set      : Loss: 1.2560 Accuracy: 0.8240 F1_score: 0.8191
```

Approach #6, use Gaussian Mixture with Auto Encoder to compute label for unlabelled data

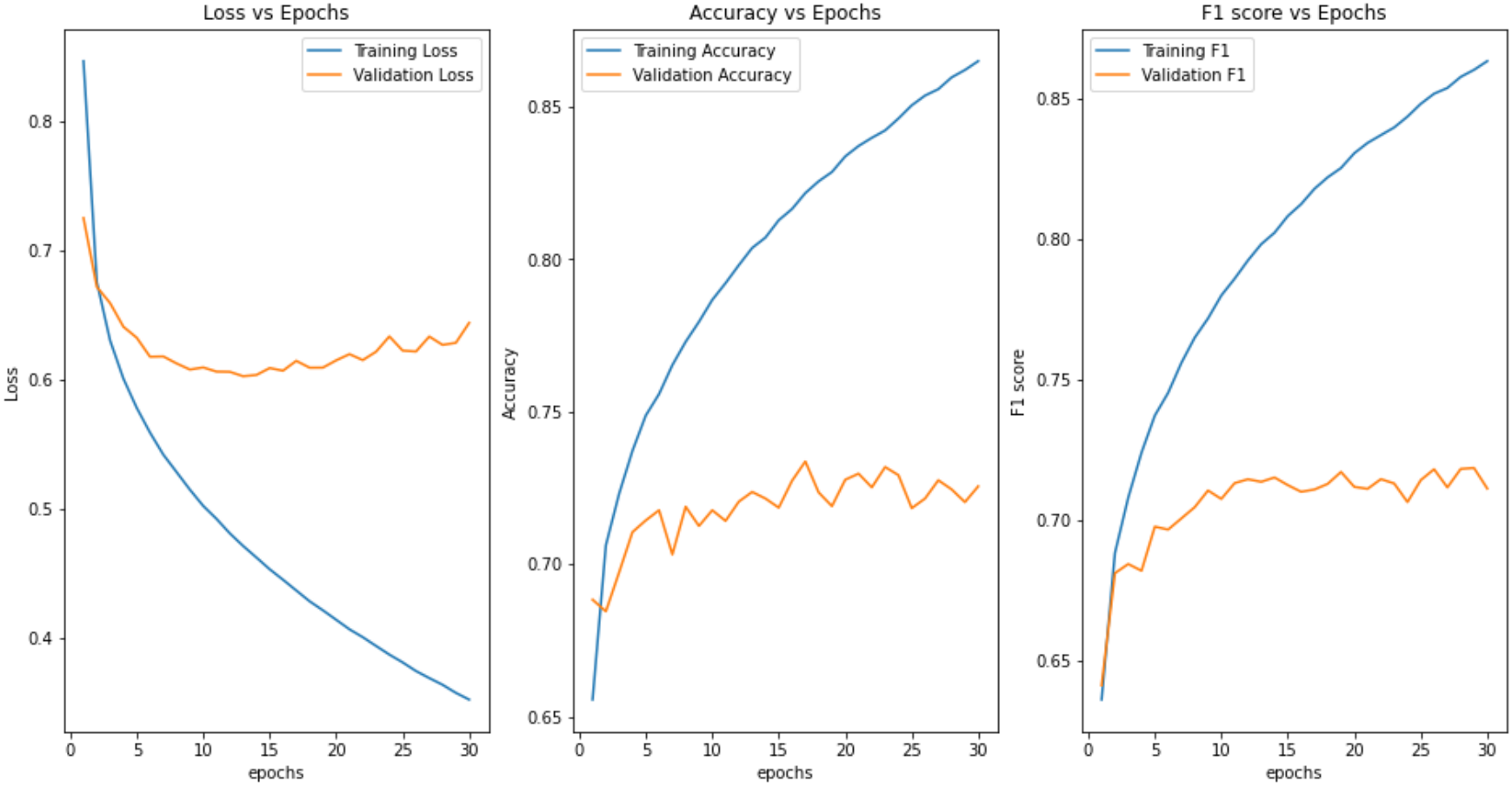
In [12]:
```
whole_flow(5, False)
```

```
Getting train and validate dataloaders for mode 5: clustering: Gaussian Mixture with Auto Encoder
Labeling unlabeled data...
Training Auto Encoder...
/ssd/jason_ssd/AdaptToNewClass/mylibs/loss.py:8: UserWarning: To copy construct from a tensor, it is recommende
d to use sourceTensor.clone().detach() or sourceTensor.clone().detach().requires_grad_(True), rather than torc
```

```
h.tensor(sourceTensor).
  log2pi = torch.log(torch.tensor(2.0 * torch.as_tensor(np.pi)))
/ssd/jason_ssd/AdaptToNewClass/mylibs/loss.py:9: UserWarning: To copy construct from a tensor, it is recommende
d to use sourceTensor.clone().detach() or sourceTensor.clone().detach().requires_grad_(True), rather than torc
h.tensor(sourceTensor).
  return torch.sum(-0.5 * ((sample - mean) ** 2.0 * torch.exp(torch.tensor(-logvar)) + logvar + log2pi), dim=ra
xis)
Epoch: 1 Loss: {'loss': 350.63329381510414}
Epoch: 2 Loss: {'loss': 320.8765460286458}
Epoch: 3 Loss: {'loss': 318.0106686035156}
Epoch: 4 Loss: {'loss': 316.85074409179686}
Epoch: 5 Loss: {'loss': 316.06245032552084}
Epoch: 6 Loss: {'loss': 315.8215092447917}
Epoch: 7 Loss: {'loss': 315.41061326497396}
Epoch: 8 Loss: {'loss': 315.20511927083334}
Epoch: 9 Loss: {'loss': 314.94296461588544}
Epoch: 10 Loss: {'loss': 314.7985603515625}
Epoch: 11 Loss: {'loss': 314.5215768066406}
Epoch: 12 Loss: {'loss': 314.42062470703127}
Epoch: 13 Loss: {'loss': 314.28416712239584}
Epoch: 14 Loss: {'loss': 314.1716448079427}
Epoch: 15 Loss: {'loss': 314.07801852213544}
Labeling accuracy: 0.3381111111111111
Start Training... Net
Epoch: 1
        Train      - Loss: 0.8468 Accuracy: 0.6556 F1_score: 0.6359
        Validation - Loss: 0.7252 Accuracy: 0.6883 F1_score: 0.6410
Epoch: 2
        Train      - Loss: 0.6757 Accuracy: 0.7062 F1_score: 0.6881
        Validation - Loss: 0.6718 Accuracy: 0.6845 F1_score: 0.6809
Epoch: 3
        Train      - Loss: 0.6302 Accuracy: 0.7231 F1_score: 0.7078
        Validation - Loss: 0.6592 Accuracy: 0.6973 F1_score: 0.6842
Epoch: 4
        Train      - Loss: 0.6007 Accuracy: 0.7371 F1_score: 0.7238
        Validation - Loss: 0.6409 Accuracy: 0.7105 F1_score: 0.6818
Epoch: 5
        Train      - Loss: 0.5780 Accuracy: 0.7487 F1_score: 0.7370
        Validation - Loss: 0.6324 Accuracy: 0.7143 F1_score: 0.6975
Epoch: 6
        Train      - Loss: 0.5588 Accuracy: 0.7558 F1_score: 0.7451
        Validation - Loss: 0.6176 Accuracy: 0.7177 F1_score: 0.6964
Epoch: 7
        Train      - Loss: 0.5418 Accuracy: 0.7653 F1_score: 0.7558
        Validation - Loss: 0.6179 Accuracy: 0.7032 F1_score: 0.7004
Epoch: 8
        Train      - Loss: 0.5282 Accuracy: 0.7729 F1_score: 0.7647
        Validation - Loss: 0.6124 Accuracy: 0.7188 F1_score: 0.7044
Epoch: 9
```

```
         Train      - Loss: 0.5146 Accuracy: 0.7794 F1_score: 0.7716
         Validation - Loss: 0.6077 Accuracy: 0.7125 F1_score: 0.7103
Epoch: 10
         Train      - Loss: 0.5020 Accuracy: 0.7866 F1_score: 0.7798
         Validation - Loss: 0.6093 Accuracy: 0.7177 F1_score: 0.7073
Epoch: 11
         Train      - Loss: 0.4920 Accuracy: 0.7921 F1_score: 0.7857
         Validation - Loss: 0.6061 Accuracy: 0.7142 F1_score: 0.7130
Epoch: 12
         Train      - Loss: 0.4808 Accuracy: 0.7980 F1_score: 0.7922
         Validation - Loss: 0.6060 Accuracy: 0.7205 F1_score: 0.7143
Epoch: 13
         Train      - Loss: 0.4710 Accuracy: 0.8037 F1_score: 0.7980
         Validation - Loss: 0.6026 Accuracy: 0.7237 F1_score: 0.7134
Epoch: 14
         Train      - Loss: 0.4620 Accuracy: 0.8071 F1_score: 0.8021
         Validation - Loss: 0.6035 Accuracy: 0.7215 F1_score: 0.7150
Epoch: 15
         Train      - Loss: 0.4529 Accuracy: 0.8128 F1_score: 0.8080
         Validation - Loss: 0.6089 Accuracy: 0.7185 F1_score: 0.7123
Epoch: 16
         Train      - Loss: 0.4448 Accuracy: 0.8164 F1_score: 0.8122
         Validation - Loss: 0.6068 Accuracy: 0.7273 F1_score: 0.7099
Epoch: 17
         Train      - Loss: 0.4364 Accuracy: 0.8217 F1_score: 0.8177
         Validation - Loss: 0.6145 Accuracy: 0.7337 F1_score: 0.7107
Epoch: 18
         Train      - Loss: 0.4281 Accuracy: 0.8255 F1_score: 0.8218
         Validation - Loss: 0.6091 Accuracy: 0.7235 F1_score: 0.7127
Epoch: 19
         Train      - Loss: 0.4211 Accuracy: 0.8286 F1_score: 0.8251
         Validation - Loss: 0.6092 Accuracy: 0.7190 F1_score: 0.7169
Epoch: 20
         Train      - Loss: 0.4137 Accuracy: 0.8337 F1_score: 0.8304
         Validation - Loss: 0.6149 Accuracy: 0.7277 F1_score: 0.7116
Epoch: 21
         Train      - Loss: 0.4063 Accuracy: 0.8371 F1_score: 0.8341
         Validation - Loss: 0.6197 Accuracy: 0.7297 F1_score: 0.7109
Epoch: 22
         Train      - Loss: 0.4002 Accuracy: 0.8398 F1_score: 0.8368
         Validation - Loss: 0.6150 Accuracy: 0.7252 F1_score: 0.7144
Epoch: 23
         Train      - Loss: 0.3934 Accuracy: 0.8422 F1_score: 0.8396
         Validation - Loss: 0.6216 Accuracy: 0.7318 F1_score: 0.7128
Epoch: 24
         Train      - Loss: 0.3867 Accuracy: 0.8461 F1_score: 0.8434
         Validation - Loss: 0.6334 Accuracy: 0.7292 F1_score: 0.7062
Epoch: 25
         Train      - Loss: 0.3807 Accuracy: 0.8504 F1_score: 0.8480
```

```
        Validation - Loss: 0.6224 Accuracy: 0.7183 F1_score: 0.7140
Epoch: 26
        Train      - Loss: 0.3740 Accuracy: 0.8536 F1_score: 0.8516
        Validation - Loss: 0.6217 Accuracy: 0.7215 F1_score: 0.7179
Epoch: 27
        Train      - Loss: 0.3685 Accuracy: 0.8557 F1_score: 0.8536
        Validation - Loss: 0.6333 Accuracy: 0.7275 F1_score: 0.7114
Epoch: 28
        Train      - Loss: 0.3633 Accuracy: 0.8596 F1_score: 0.8575
        Validation - Loss: 0.6268 Accuracy: 0.7245 F1_score: 0.7180
Epoch: 29
        Train      - Loss: 0.3571 Accuracy: 0.8620 F1_score: 0.8601
        Validation - Loss: 0.6284 Accuracy: 0.7203 F1_score: 0.7184
Epoch: 30
        Train      - Loss: 0.3519 Accuracy: 0.8649 F1_score: 0.8632
        Validation - Loss: 0.6439 Accuracy: 0.7255 F1_score: 0.7110
<Figure size 432x288 with 0 Axes>
```



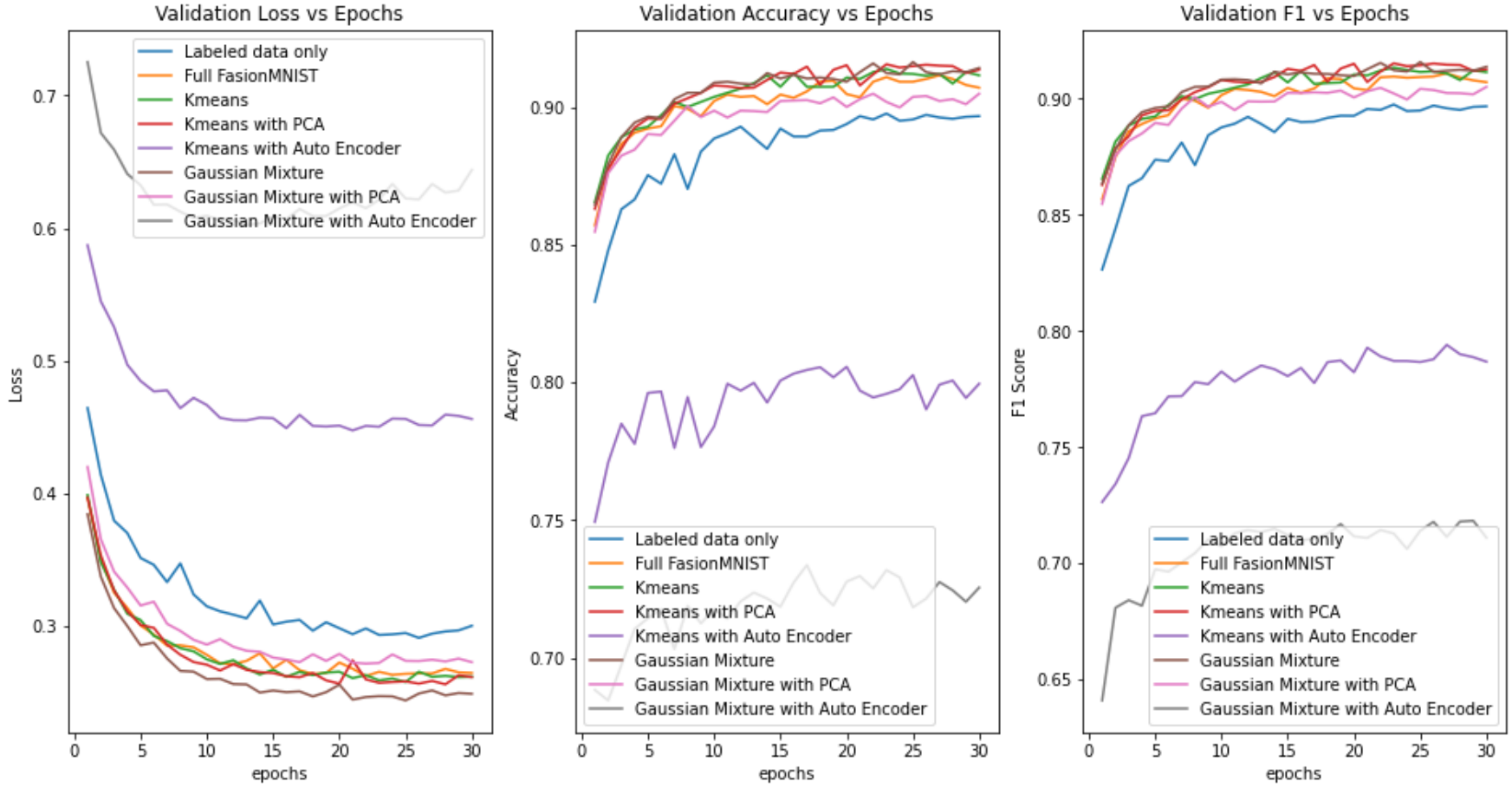

```
Loading model at epoch 29 for best validation f1
```

```
Preparing test loaders
Result on labelled test set   : Loss: 0.3594 Accuracy: 0.8769 F1_score: 0.8816
Result on unlabelled test set: Loss: 1.3512 Accuracy: 0.3557 F1_score: 0.3060
Result on full test set       : Loss: 0.6554 Accuracy: 0.7205 F1_score: 0.7052
```

Summary

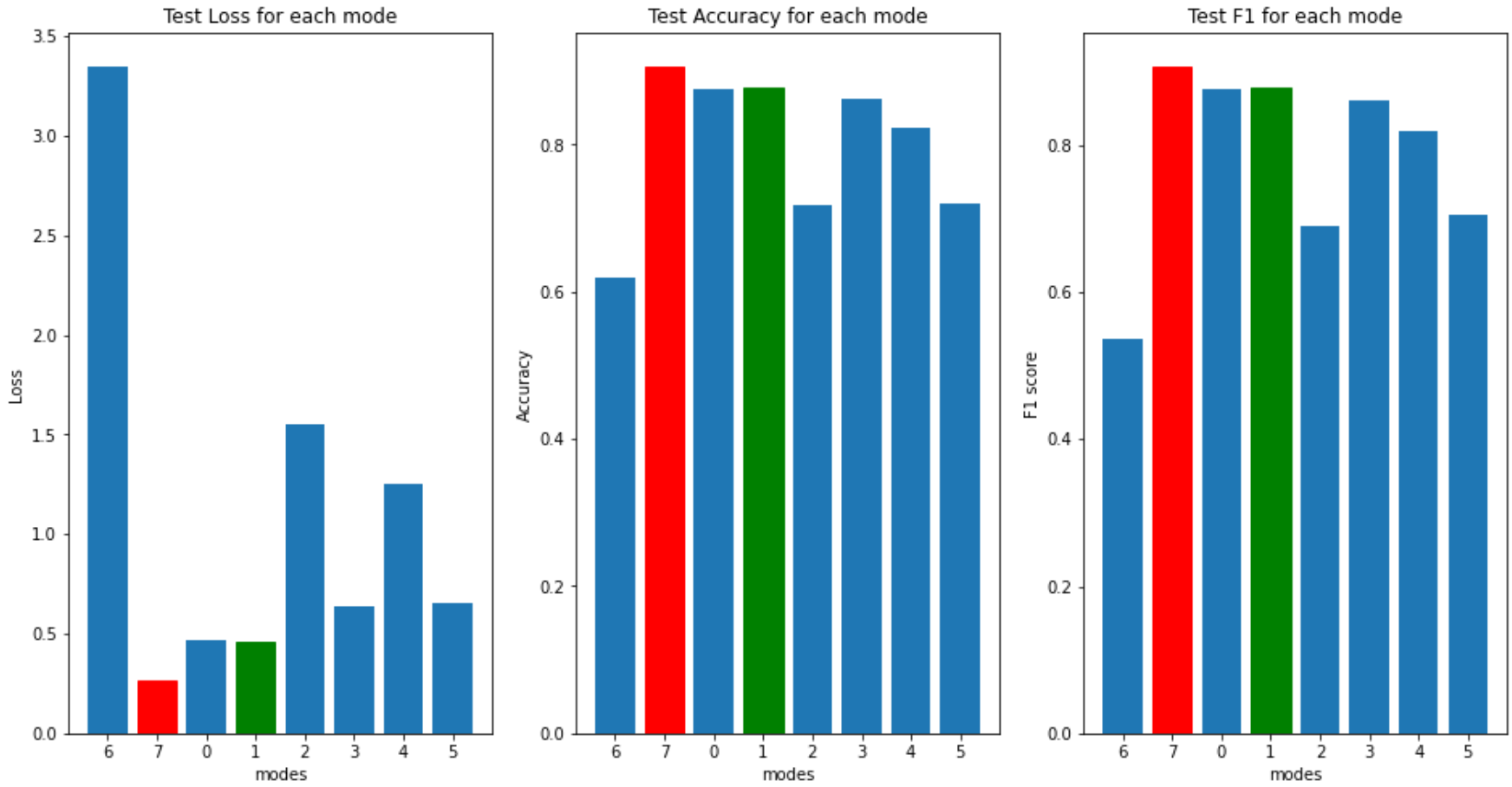
In [13]:
```
report_summary(mode_metrics, mode_description_short)
```

```
<Figure size 432x288 with 0 Axes>
```



In [18]:
```
report_test_summary(mode_test_metrics, mode_description_short)
```

```
<Figure size 432x288 with 0 Axes>
```




```
Mode Interpretations:
    6: Labeled data only
    7: Full FasionMNIST
    0: Kmeans
    1: Kmeans with PCA
    2: Kmeans with Auto Encoder
    3: Gaussian Mixture
    4: Gaussian Mixture with PCA
    5: Gaussian Mixture with Auto Encoder
Color Interpretations:
    Red  : Best result over all runs
    Green: Best result over all approachs (excluding baselines)
```

In [27]:
```
clustering_acc()
```

```
- Unsupervised Clusteriing on the full test set:
Labeling unlabeled data...
Labeling accuracy: 0.5603
```

```
- Unsupervised Clusteriing on the labelled test set:
Labeling unlabeled data...
Labeling accuracy: 0.551
- Unsupervised Clusteriing on the unlabelled test set:
Labeling unlabeled data...
Labeling accuracy: 0.8876666666666667
```

Method	Test Loss on Labeled Data	Test Accuracy on Labeled Data	Test F1 on Labeled Data	Test Loss on Unlabeled Data	Test Accuracy on Unlabeled Data	Test F1 on Unlabeled Data	Test Loss on All Data	Test Accuracy on All Data	Test F1 on All Data
Baseline 1 - Only use Labeled data	0.3275	0.8856	0.8856	10.4008	0.0000	0.0000	3.3507	0.6199	0.5362
KMeans	0.3499	0.8784	0.8829	0.7429	0.8683	0.8772	0.4662	0.8754	0.8763
KMeans with PCA	0.3549	0.8773	0.8808	0.7100	0.8773	0.8850	0.4614	0.8773	0.8769
KMeans with VAE	0.3618	0.8753	0.8790	4.3273	0.3537	0.2550	1.5523	0.7188	0.6891
Gaussian Mixture	0.3434	0.8786	0.8819	1.3362	0.8250	0.8280	0.6390	0.8625	0.8614
Gaussian Mixture with PCA	0.3369	0.8830	0.8866	3.3661	0.6863	0.6731	1.2560	0.8240	0.8191
Gaussian Mixture with VAE	0.3594	0.8769	0.8816	1.3512	0.3557	0.3060	0.6554	0.7205	0.7052
Baseline 2 - Full FashionMNIST data	0.3431	0.8801	0.8843	0.0965	0.9697	0.9762	0.2685	0.9070	0.9073
Unsupervised Clustering	N/A	0.551	N/A	N/A	0.8876	N/A	N/A	0.5603	N/A

Part 2: ResNet model training

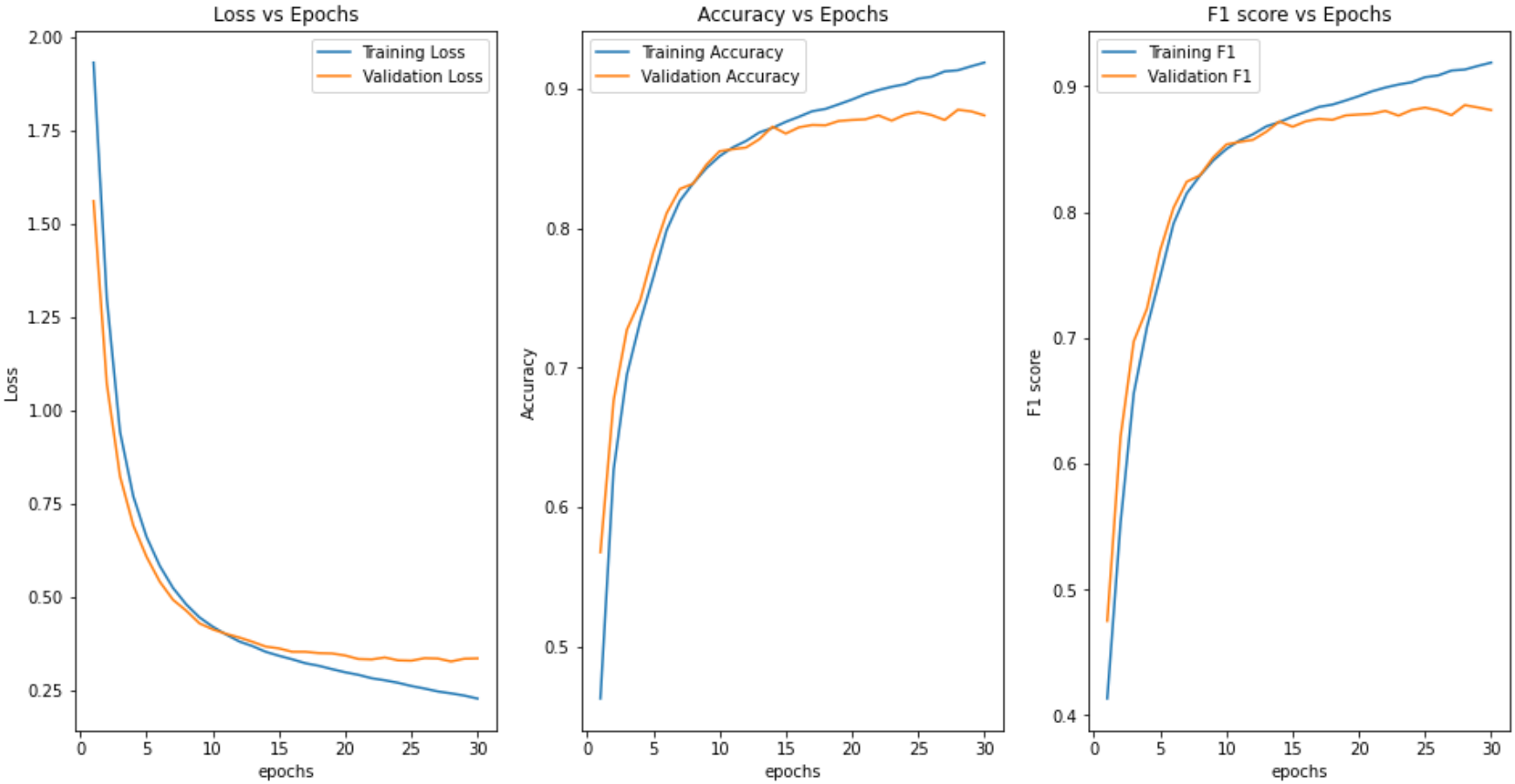
in this part, the Neural Network is replaced from our CNN to pre-trained ResNet. The purpose is to compare between neural networks.

Baseline #1, use only first K-class labelled data to train

In [5]:
```
whole_flow(6, True)
```

```
Getting train and validate dataloaders for mode 6: use only labeled data
Start Training... ResNet
Epoch: 1
        Train      - Loss: 1.9306 Accuracy: 0.4626 F1_score: 0.4133
        Validation - Loss: 1.5597 Accuracy: 0.5676 F1_score: 0.4750
Epoch: 2
        Train      - Loss: 1.2985 Accuracy: 0.6270 F1_score: 0.5527
        Validation - Loss: 1.0706 Accuracy: 0.6767 F1_score: 0.6212
Epoch: 3
        Train      - Loss: 0.9415 Accuracy: 0.6951 F1_score: 0.6558
        Validation - Loss: 0.8223 Accuracy: 0.7274 F1_score: 0.6972
Epoch: 4
        Train      - Loss: 0.7693 Accuracy: 0.7330 F1_score: 0.7085
        Validation - Loss: 0.6914 Accuracy: 0.7483 F1_score: 0.7232
Epoch: 5
        Train      - Loss: 0.6605 Accuracy: 0.7651 F1_score: 0.7488
        Validation - Loss: 0.6072 Accuracy: 0.7829 F1_score: 0.7698
Epoch: 6
        Train      - Loss: 0.5830 Accuracy: 0.7987 F1_score: 0.7907
        Validation - Loss: 0.5405 Accuracy: 0.8110 F1_score: 0.8033
Epoch: 7
        Train      - Loss: 0.5238 Accuracy: 0.8198 F1_score: 0.8149
        Validation - Loss: 0.4920 Accuracy: 0.8283 F1_score: 0.8241
Epoch: 8
        Train      - Loss: 0.4796 Accuracy: 0.8323 F1_score: 0.8289
        Validation - Loss: 0.4629 Accuracy: 0.8321 F1_score: 0.8292
Epoch: 9
        Train      - Loss: 0.4449 Accuracy: 0.8433 F1_score: 0.8410
        Validation - Loss: 0.4289 Accuracy: 0.8457 F1_score: 0.8434
Epoch: 10
        Train      - Loss: 0.4199 Accuracy: 0.8519 F1_score: 0.8502
        Validation - Loss: 0.4129 Accuracy: 0.8555 F1_score: 0.8538
Epoch: 11
        Train      - Loss: 0.3992 Accuracy: 0.8582 F1_score: 0.8570
        Validation - Loss: 0.4008 Accuracy: 0.8569 F1_score: 0.8557
Epoch: 12
        Train      - Loss: 0.3804 Accuracy: 0.8628 F1_score: 0.8618
        Validation - Loss: 0.3910 Accuracy: 0.8581 F1_score: 0.8574
Epoch: 13
        Train      - Loss: 0.3679 Accuracy: 0.8689 F1_score: 0.8681
        Validation - Loss: 0.3795 Accuracy: 0.8640 F1_score: 0.8638
Epoch: 14
        Train      - Loss: 0.3527 Accuracy: 0.8722 F1_score: 0.8715
        Validation - Loss: 0.3667 Accuracy: 0.8731 F1_score: 0.8720
Epoch: 15
        Train      - Loss: 0.3421 Accuracy: 0.8765 F1_score: 0.8759
```

```
        Validation - Loss: 0.3617 Accuracy: 0.8681 F1_score: 0.8678
Epoch: 16
        Train      - Loss: 0.3326 Accuracy: 0.8802 F1_score: 0.8797
        Validation - Loss: 0.3528 Accuracy: 0.8726 F1_score: 0.8722
Epoch: 17
        Train      - Loss: 0.3221 Accuracy: 0.8842 F1_score: 0.8836
        Validation - Loss: 0.3527 Accuracy: 0.8743 F1_score: 0.8741
Epoch: 18
        Train      - Loss: 0.3153 Accuracy: 0.8858 F1_score: 0.8854
        Validation - Loss: 0.3493 Accuracy: 0.8740 F1_score: 0.8733
Epoch: 19
        Train      - Loss: 0.3063 Accuracy: 0.8891 F1_score: 0.8887
        Validation - Loss: 0.3483 Accuracy: 0.8771 F1_score: 0.8768
Epoch: 20
        Train      - Loss: 0.2979 Accuracy: 0.8925 F1_score: 0.8922
        Validation - Loss: 0.3429 Accuracy: 0.8779 F1_score: 0.8775
Epoch: 21
        Train      - Loss: 0.2910 Accuracy: 0.8964 F1_score: 0.8960
        Validation - Loss: 0.3334 Accuracy: 0.8783 F1_score: 0.8781
Epoch: 22
        Train      - Loss: 0.2819 Accuracy: 0.8993 F1_score: 0.8990
        Validation - Loss: 0.3323 Accuracy: 0.8812 F1_score: 0.8804
Epoch: 23
        Train      - Loss: 0.2763 Accuracy: 0.9017 F1_score: 0.9014
        Validation - Loss: 0.3376 Accuracy: 0.8774 F1_score: 0.8766
Epoch: 24
        Train      - Loss: 0.2697 Accuracy: 0.9035 F1_score: 0.9032
        Validation - Loss: 0.3298 Accuracy: 0.8817 F1_score: 0.8811
Epoch: 25
        Train      - Loss: 0.2611 Accuracy: 0.9075 F1_score: 0.9072
        Validation - Loss: 0.3289 Accuracy: 0.8836 F1_score: 0.8830
Epoch: 26
        Train      - Loss: 0.2542 Accuracy: 0.9089 F1_score: 0.9086
        Validation - Loss: 0.3356 Accuracy: 0.8814 F1_score: 0.8808
Epoch: 27
        Train      - Loss: 0.2465 Accuracy: 0.9128 F1_score: 0.9125
        Validation - Loss: 0.3346 Accuracy: 0.8779 F1_score: 0.8770
Epoch: 28
        Train      - Loss: 0.2414 Accuracy: 0.9135 F1_score: 0.9133
        Validation - Loss: 0.3268 Accuracy: 0.8852 F1_score: 0.8850
Epoch: 29
        Train      - Loss: 0.2355 Accuracy: 0.9164 F1_score: 0.9162
        Validation - Loss: 0.3341 Accuracy: 0.8840 F1_score: 0.8831
Epoch: 30
        Train      - Loss: 0.2274 Accuracy: 0.9191 F1_score: 0.9188
        Validation - Loss: 0.3350 Accuracy: 0.8812 F1_score: 0.8811
<Figure size 432x288 with 0 Axes>
```




```
Loading model at epoch 28 for best validation f1
Preparing test loaders
Result on labelled test set   : Loss: 0.3572 Accuracy: 0.8726 F1_score: 0.8729
Result on unlabelled test set: Loss: 9.4141 Accuracy: 0.0000 F1_score: 0.0000
Result on full test set       : Loss: 3.0787 Accuracy: 0.6108 F1_score: 0.5264
```

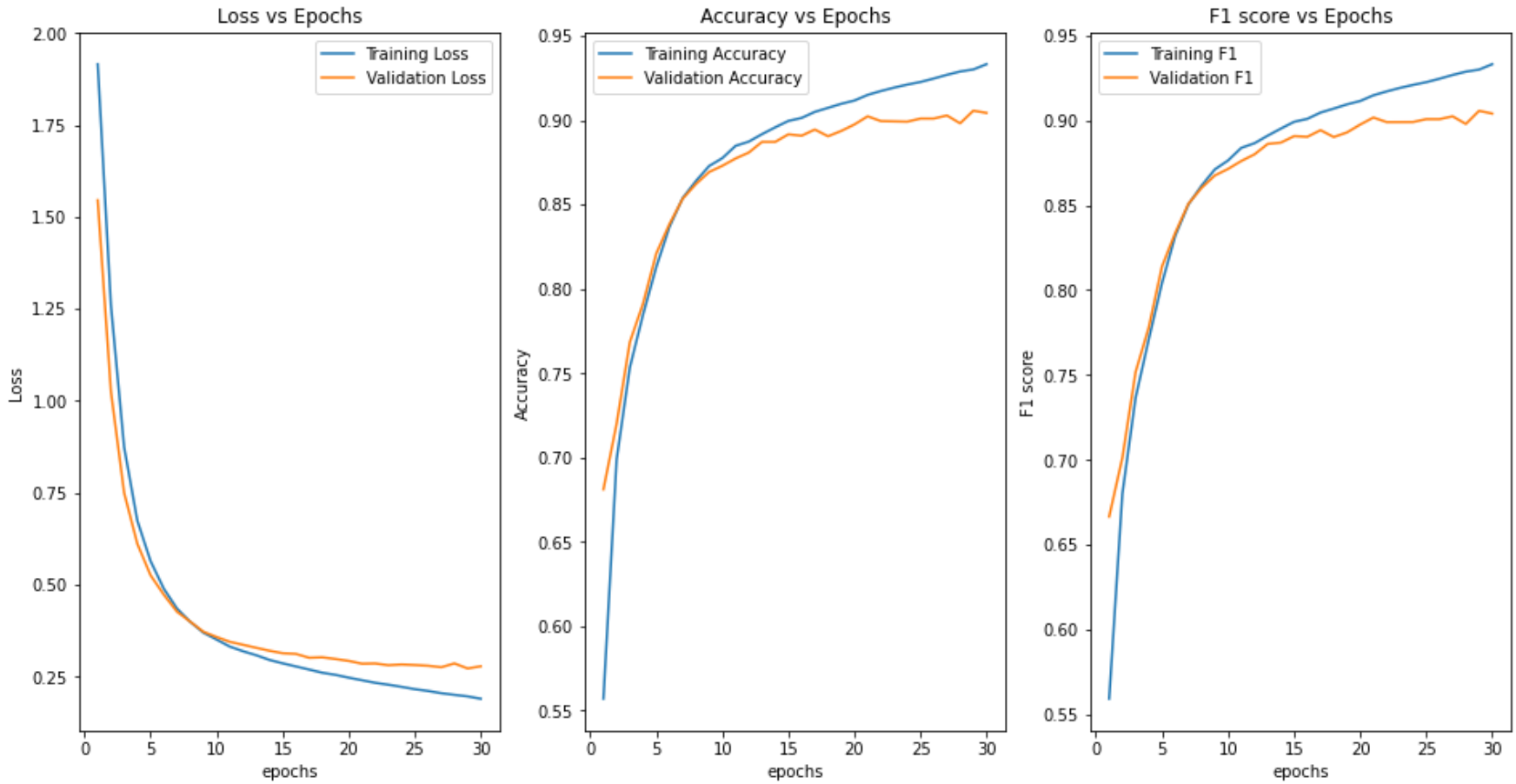
Baseline #2, use full Fashion-MNIST dataset to train

In [6]:
```
whole_flow(7, True)
```

```
Getting train and validate dataloaders for mode 7: use full FasionMNIST data
Start Training... ResNet
Epoch: 1
        Train      - Loss: 1.9156 Accuracy: 0.5571 F1_score: 0.5593
        Validation - Loss: 1.5457 Accuracy: 0.6812 F1_score: 0.6666
Epoch: 2
        Train      - Loss: 1.2603 Accuracy: 0.6990 F1_score: 0.6801
```

```
         Validation - Loss: 1.0255 Accuracy: 0.7202 F1_score: 0.7013
Epoch: 3
         Train      - Loss: 0.8744 Accuracy: 0.7535 F1_score: 0.7366
         Validation - Loss: 0.7502 Accuracy: 0.7685 F1_score: 0.7517
Epoch: 4
         Train      - Loss: 0.6742 Accuracy: 0.7848 F1_score: 0.7711
         Validation - Loss: 0.6113 Accuracy: 0.7913 F1_score: 0.7781
Epoch: 5
         Train      - Loss: 0.5643 Accuracy: 0.8130 F1_score: 0.8042
         Validation - Loss: 0.5257 Accuracy: 0.8210 F1_score: 0.8140
Epoch: 6
         Train      - Loss: 0.4893 Accuracy: 0.8368 F1_score: 0.8320
         Validation - Loss: 0.4731 Accuracy: 0.8385 F1_score: 0.8337
Epoch: 7
         Train      - Loss: 0.4351 Accuracy: 0.8539 F1_score: 0.8505
         Validation - Loss: 0.4266 Accuracy: 0.8535 F1_score: 0.8509
Epoch: 8
         Train      - Loss: 0.3992 Accuracy: 0.8640 F1_score: 0.8616
         Validation - Loss: 0.3995 Accuracy: 0.8622 F1_score: 0.8603
Epoch: 9
         Train      - Loss: 0.3695 Accuracy: 0.8729 F1_score: 0.8711
         Validation - Loss: 0.3720 Accuracy: 0.8693 F1_score: 0.8676
Epoch: 10
         Train      - Loss: 0.3507 Accuracy: 0.8775 F1_score: 0.8764
         Validation - Loss: 0.3576 Accuracy: 0.8730 F1_score: 0.8714
Epoch: 11
         Train      - Loss: 0.3316 Accuracy: 0.8848 F1_score: 0.8838
         Validation - Loss: 0.3447 Accuracy: 0.8773 F1_score: 0.8761
Epoch: 12
         Train      - Loss: 0.3191 Accuracy: 0.8873 F1_score: 0.8865
         Validation - Loss: 0.3366 Accuracy: 0.8808 F1_score: 0.8800
Epoch: 13
         Train      - Loss: 0.3082 Accuracy: 0.8917 F1_score: 0.8910
         Validation - Loss: 0.3284 Accuracy: 0.8872 F1_score: 0.8861
Epoch: 14
         Train      - Loss: 0.2954 Accuracy: 0.8958 F1_score: 0.8952
         Validation - Loss: 0.3204 Accuracy: 0.8872 F1_score: 0.8869
Epoch: 15
         Train      - Loss: 0.2862 Accuracy: 0.8996 F1_score: 0.8991
         Validation - Loss: 0.3137 Accuracy: 0.8917 F1_score: 0.8907
Epoch: 16
         Train      - Loss: 0.2777 Accuracy: 0.9014 F1_score: 0.9009
         Validation - Loss: 0.3118 Accuracy: 0.8908 F1_score: 0.8903
Epoch: 17
         Train      - Loss: 0.2692 Accuracy: 0.9050 F1_score: 0.9046
         Validation - Loss: 0.3013 Accuracy: 0.8945 F1_score: 0.8942
Epoch: 18
         Train      - Loss: 0.2606 Accuracy: 0.9074 F1_score: 0.9069
         Validation - Loss: 0.3025 Accuracy: 0.8905 F1_score: 0.8902
```

```
Epoch: 19
        Train      - Loss: 0.2548 Accuracy: 0.9097 F1_score: 0.9094
        Validation - Loss: 0.2979 Accuracy: 0.8937 F1_score: 0.8929
Epoch: 20
        Train      - Loss: 0.2471 Accuracy: 0.9117 F1_score: 0.9114
        Validation - Loss: 0.2927 Accuracy: 0.8975 F1_score: 0.8975
Epoch: 21
        Train      - Loss: 0.2401 Accuracy: 0.9150 F1_score: 0.9148
        Validation - Loss: 0.2852 Accuracy: 0.9023 F1_score: 0.9017
Epoch: 22
        Train      - Loss: 0.2332 Accuracy: 0.9173 F1_score: 0.9170
        Validation - Loss: 0.2858 Accuracy: 0.8995 F1_score: 0.8990
Epoch: 23
        Train      - Loss: 0.2280 Accuracy: 0.9194 F1_score: 0.9191
        Validation - Loss: 0.2810 Accuracy: 0.8993 F1_score: 0.8990
Epoch: 24
        Train      - Loss: 0.2222 Accuracy: 0.9211 F1_score: 0.9209
        Validation - Loss: 0.2830 Accuracy: 0.8992 F1_score: 0.8990
Epoch: 25
        Train      - Loss: 0.2158 Accuracy: 0.9227 F1_score: 0.9225
        Validation - Loss: 0.2815 Accuracy: 0.9010 F1_score: 0.9007
Epoch: 26
        Train      - Loss: 0.2110 Accuracy: 0.9247 F1_score: 0.9245
        Validation - Loss: 0.2796 Accuracy: 0.9010 F1_score: 0.9007
Epoch: 27
        Train      - Loss: 0.2050 Accuracy: 0.9269 F1_score: 0.9267
        Validation - Loss: 0.2755 Accuracy: 0.9028 F1_score: 0.9024
Epoch: 28
        Train      - Loss: 0.2005 Accuracy: 0.9289 F1_score: 0.9287
        Validation - Loss: 0.2859 Accuracy: 0.8982 F1_score: 0.8979
Epoch: 29
        Train      - Loss: 0.1962 Accuracy: 0.9301 F1_score: 0.9299
        Validation - Loss: 0.2722 Accuracy: 0.9057 F1_score: 0.9056
Epoch: 30
        Train      - Loss: 0.1898 Accuracy: 0.9332 F1_score: 0.9331
        Validation - Loss: 0.2781 Accuracy: 0.9043 F1_score: 0.9040
<Figure size 432x288 with 0 Axes>
```




```
Loading model at epoch 29 for best validation f1
Preparing test loaders
Result on labelled test set  : Loss: 0.3777 Accuracy: 0.8690 F1_score: 0.8719
Result on unlabelled test set: Loss: 0.0975 Accuracy: 0.9683 F1_score: 0.9741
Result on full test set      : Loss: 0.2934 Accuracy: 0.8988 F1_score: 0.8984
```

Label new categories by clustering

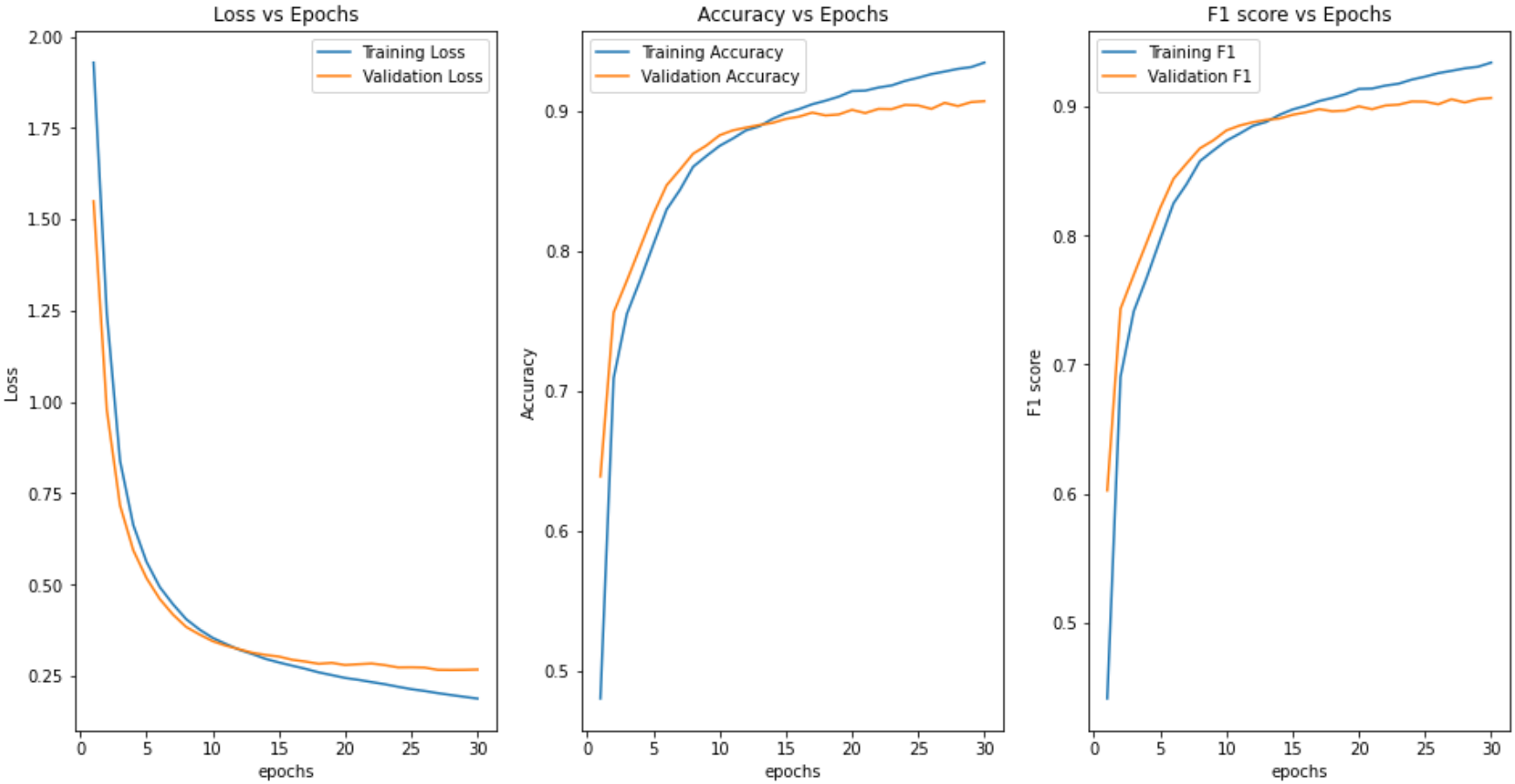
- label the dataset under different clustering methods
- compare acuracy on the test set

Approach #1, use KMeans to compute label for unlabelled data

In [7]:
```
whole_flow(0, True)
```

```
Getting train and validate dataloaders for mode 0: clustering: kmeans
Labeling unlabeled data...
Labeling accuracy: 0.8799444444444444
Start Training... ResNet
Epoch: 1
        Train      - Loss: 1.9291 Accuracy: 0.4805 F1_score: 0.4409
        Validation - Loss: 1.5497 Accuracy: 0.6390 F1_score: 0.6023
Epoch: 2
        Train      - Loss: 1.2419 Accuracy: 0.7092 F1_score: 0.6905
        Validation - Loss: 0.9784 Accuracy: 0.7560 F1_score: 0.7432
Epoch: 3
        Train      - Loss: 0.8392 Accuracy: 0.7551 F1_score: 0.7412
        Validation - Loss: 0.7171 Accuracy: 0.7788 F1_score: 0.7695
Epoch: 4
        Train      - Loss: 0.6636 Accuracy: 0.7791 F1_score: 0.7680
        Validation - Loss: 0.5943 Accuracy: 0.8025 F1_score: 0.7953
Epoch: 5
        Train      - Loss: 0.5626 Accuracy: 0.8046 F1_score: 0.7969
        Validation - Loss: 0.5180 Accuracy: 0.8263 F1_score: 0.8216
Epoch: 6
        Train      - Loss: 0.4937 Accuracy: 0.8295 F1_score: 0.8250
        Validation - Loss: 0.4609 Accuracy: 0.8468 F1_score: 0.8440
Epoch: 7
        Train      - Loss: 0.4466 Accuracy: 0.8435 F1_score: 0.8401
        Validation - Loss: 0.4183 Accuracy: 0.8578 F1_score: 0.8561
Epoch: 8
        Train      - Loss: 0.4057 Accuracy: 0.8600 F1_score: 0.8578
        Validation - Loss: 0.3844 Accuracy: 0.8693 F1_score: 0.8677
Epoch: 9
        Train      - Loss: 0.3775 Accuracy: 0.8677 F1_score: 0.8659
        Validation - Loss: 0.3638 Accuracy: 0.8752 F1_score: 0.8738
Epoch: 10
        Train      - Loss: 0.3545 Accuracy: 0.8749 F1_score: 0.8736
        Validation - Loss: 0.3455 Accuracy: 0.8825 F1_score: 0.8816
Epoch: 11
        Train      - Loss: 0.3378 Accuracy: 0.8802 F1_score: 0.8791
        Validation - Loss: 0.3332 Accuracy: 0.8860 F1_score: 0.8854
Epoch: 12
        Train      - Loss: 0.3221 Accuracy: 0.8860 F1_score: 0.8852
        Validation - Loss: 0.3235 Accuracy: 0.8880 F1_score: 0.8880
Epoch: 13
        Train      - Loss: 0.3103 Accuracy: 0.8889 F1_score: 0.8882
        Validation - Loss: 0.3139 Accuracy: 0.8898 F1_score: 0.8898
Epoch: 14
        Train      - Loss: 0.2969 Accuracy: 0.8942 F1_score: 0.8937
        Validation - Loss: 0.3078 Accuracy: 0.8913 F1_score: 0.8908
Epoch: 15
        Train      - Loss: 0.2873 Accuracy: 0.8983 F1_score: 0.8979
```

```
        Validation - Loss: 0.3034 Accuracy: 0.8942 F1_score: 0.8938
Epoch: 16
        Train      - Loss: 0.2785 Accuracy: 0.9012 F1_score: 0.9007
        Validation - Loss: 0.2947 Accuracy: 0.8958 F1_score: 0.8956
Epoch: 17
        Train      - Loss: 0.2699 Accuracy: 0.9046 F1_score: 0.9044
        Validation - Loss: 0.2897 Accuracy: 0.8987 F1_score: 0.8981
Epoch: 18
        Train      - Loss: 0.2603 Accuracy: 0.9072 F1_score: 0.9068
        Validation - Loss: 0.2837 Accuracy: 0.8967 F1_score: 0.8965
Epoch: 19
        Train      - Loss: 0.2526 Accuracy: 0.9101 F1_score: 0.9098
        Validation - Loss: 0.2862 Accuracy: 0.8973 F1_score: 0.8970
Epoch: 20
        Train      - Loss: 0.2449 Accuracy: 0.9140 F1_score: 0.9137
        Validation - Loss: 0.2803 Accuracy: 0.9007 F1_score: 0.9003
Epoch: 21
        Train      - Loss: 0.2397 Accuracy: 0.9144 F1_score: 0.9141
        Validation - Loss: 0.2825 Accuracy: 0.8983 F1_score: 0.8981
Epoch: 22
        Train      - Loss: 0.2336 Accuracy: 0.9166 F1_score: 0.9164
        Validation - Loss: 0.2846 Accuracy: 0.9013 F1_score: 0.9009
Epoch: 23
        Train      - Loss: 0.2278 Accuracy: 0.9181 F1_score: 0.9178
        Validation - Loss: 0.2800 Accuracy: 0.9012 F1_score: 0.9016
Epoch: 24
        Train      - Loss: 0.2207 Accuracy: 0.9213 F1_score: 0.9210
        Validation - Loss: 0.2737 Accuracy: 0.9042 F1_score: 0.9040
Epoch: 25
        Train      - Loss: 0.2142 Accuracy: 0.9236 F1_score: 0.9233
        Validation - Loss: 0.2740 Accuracy: 0.9038 F1_score: 0.9039
Epoch: 26
        Train      - Loss: 0.2092 Accuracy: 0.9262 F1_score: 0.9260
        Validation - Loss: 0.2732 Accuracy: 0.9013 F1_score: 0.9019
Epoch: 27
        Train      - Loss: 0.2033 Accuracy: 0.9280 F1_score: 0.9278
        Validation - Loss: 0.2670 Accuracy: 0.9057 F1_score: 0.9058
Epoch: 28
        Train      - Loss: 0.1982 Accuracy: 0.9299 F1_score: 0.9297
        Validation - Loss: 0.2666 Accuracy: 0.9033 F1_score: 0.9033
Epoch: 29
        Train      - Loss: 0.1932 Accuracy: 0.9311 F1_score: 0.9310
        Validation - Loss: 0.2670 Accuracy: 0.9062 F1_score: 0.9059
Epoch: 30
        Train      - Loss: 0.1884 Accuracy: 0.9344 F1_score: 0.9342
        Validation - Loss: 0.2679 Accuracy: 0.9068 F1_score: 0.9068
<Figure size 432x288 with 0 Axes>
```




```
Loading model at epoch 30 for best validation f1
Preparing test loaders
Result on labelled test set  : Loss: 0.3637 Accuracy: 0.8716 F1_score: 0.8742
Result on unlabelled test set: Loss: 0.6130 Accuracy: 0.8737 F1_score: 0.8823
Result on full test set      : Loss: 0.4424 Accuracy: 0.8722 F1_score: 0.8729
```

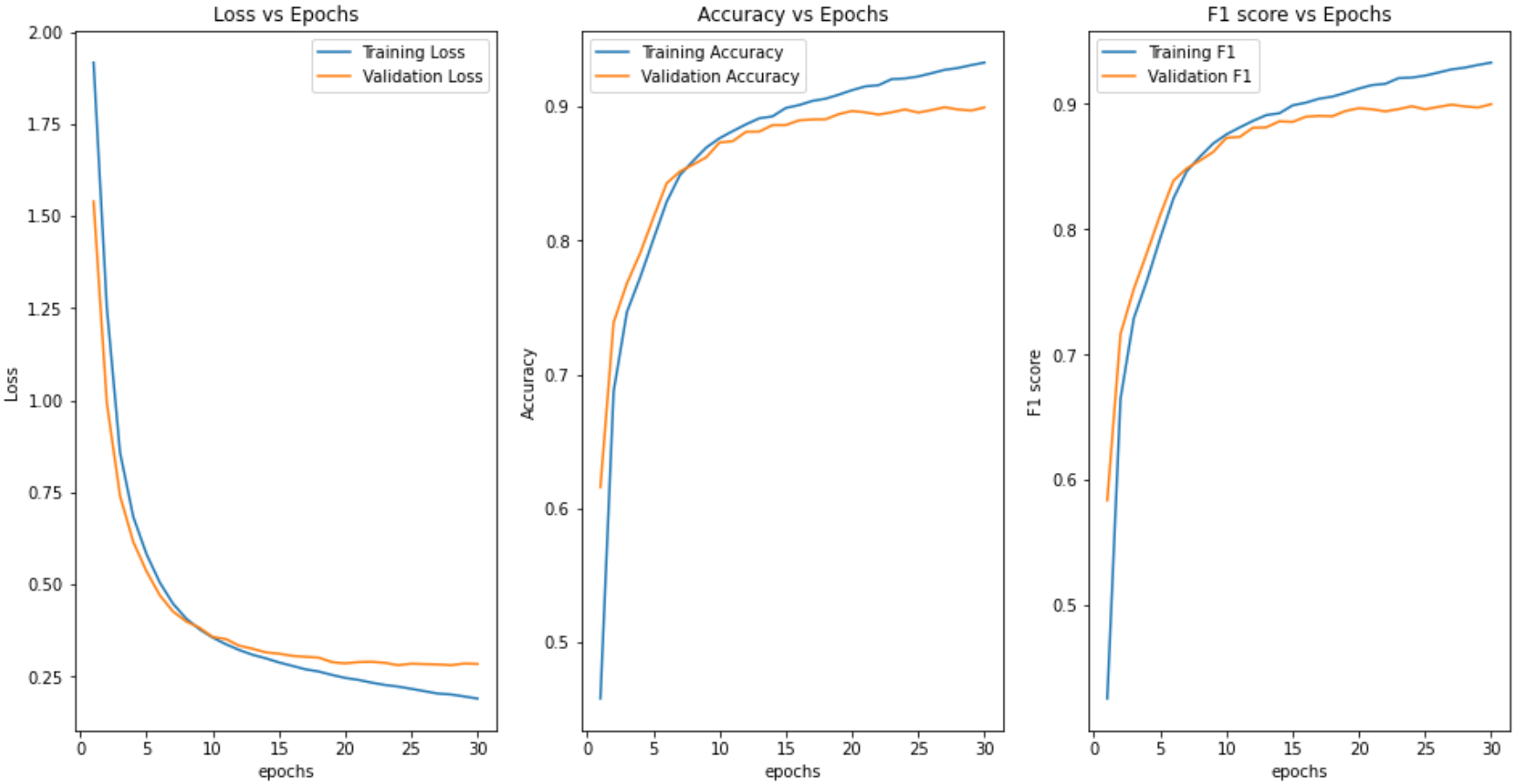
Approach #2, use KMeans with PCA to compute label for unlabelled data

In [8]:
```
whole_flow(1, True)
```

```
Getting train and validate dataloaders for mode 1: clustering: kmeans with PCA
Labeling unlabeled data...
Labeling accuracy: 0.8811111111111111
Start Training... ResNet
Epoch: 1
        Train      - Loss: 1.9170 Accuracy: 0.4582 F1_score: 0.4258
        Validation - Loss: 1.5401 Accuracy: 0.6160 F1_score: 0.5838
```

```
Epoch: 2
        Train      - Loss: 1.2525 Accuracy: 0.6877 F1_score: 0.6648
        Validation - Loss: 0.9935 Accuracy: 0.7392 F1_score: 0.7165
Epoch: 3
        Train      - Loss: 0.8581 Accuracy: 0.7468 F1_score: 0.7288
        Validation - Loss: 0.7403 Accuracy: 0.7682 F1_score: 0.7522
Epoch: 4
        Train      - Loss: 0.6837 Accuracy: 0.7726 F1_score: 0.7592
        Validation - Loss: 0.6155 Accuracy: 0.7907 F1_score: 0.7813
Epoch: 5
        Train      - Loss: 0.5823 Accuracy: 0.8010 F1_score: 0.7929
        Validation - Loss: 0.5355 Accuracy: 0.8172 F1_score: 0.8115
Epoch: 6
        Train      - Loss: 0.5054 Accuracy: 0.8290 F1_score: 0.8246
        Validation - Loss: 0.4708 Accuracy: 0.8428 F1_score: 0.8387
Epoch: 7
        Train      - Loss: 0.4478 Accuracy: 0.8489 F1_score: 0.8462
        Validation - Loss: 0.4264 Accuracy: 0.8517 F1_score: 0.8485
Epoch: 8
        Train      - Loss: 0.4071 Accuracy: 0.8596 F1_score: 0.8579
        Validation - Loss: 0.3995 Accuracy: 0.8568 F1_score: 0.8548
Epoch: 9
        Train      - Loss: 0.3784 Accuracy: 0.8696 F1_score: 0.8684
        Validation - Loss: 0.3825 Accuracy: 0.8623 F1_score: 0.8616
Epoch: 10
        Train      - Loss: 0.3556 Accuracy: 0.8765 F1_score: 0.8756
        Validation - Loss: 0.3574 Accuracy: 0.8733 F1_score: 0.8729
Epoch: 11
        Train      - Loss: 0.3376 Accuracy: 0.8819 F1_score: 0.8812
        Validation - Loss: 0.3515 Accuracy: 0.8743 F1_score: 0.8736
Epoch: 12
        Train      - Loss: 0.3220 Accuracy: 0.8869 F1_score: 0.8864
        Validation - Loss: 0.3336 Accuracy: 0.8813 F1_score: 0.8809
Epoch: 13
        Train      - Loss: 0.3091 Accuracy: 0.8914 F1_score: 0.8910
        Validation - Loss: 0.3254 Accuracy: 0.8815 F1_score: 0.8813
Epoch: 14
        Train      - Loss: 0.2995 Accuracy: 0.8929 F1_score: 0.8925
        Validation - Loss: 0.3157 Accuracy: 0.8863 F1_score: 0.8861
Epoch: 15
        Train      - Loss: 0.2883 Accuracy: 0.8991 F1_score: 0.8988
        Validation - Loss: 0.3121 Accuracy: 0.8863 F1_score: 0.8856
Epoch: 16
        Train      - Loss: 0.2790 Accuracy: 0.9013 F1_score: 0.9009
        Validation - Loss: 0.3062 Accuracy: 0.8898 F1_score: 0.8898
Epoch: 17
        Train      - Loss: 0.2695 Accuracy: 0.9043 F1_score: 0.9041
        Validation - Loss: 0.3035 Accuracy: 0.8905 F1_score: 0.8904
Epoch: 18
```

```
         Train      - Loss: 0.2637 Accuracy: 0.9060 F1_score: 0.9058
         Validation - Loss: 0.3014 Accuracy: 0.8907 F1_score: 0.8901
Epoch: 19
         Train      - Loss: 0.2544 Accuracy: 0.9090 F1_score: 0.9088
         Validation - Loss: 0.2892 Accuracy: 0.8945 F1_score: 0.8943
Epoch: 20
         Train      - Loss: 0.2463 Accuracy: 0.9124 F1_score: 0.9122
         Validation - Loss: 0.2857 Accuracy: 0.8968 F1_score: 0.8965
Epoch: 21
         Train      - Loss: 0.2408 Accuracy: 0.9152 F1_score: 0.9150
         Validation - Loss: 0.2892 Accuracy: 0.8958 F1_score: 0.8957
Epoch: 22
         Train      - Loss: 0.2334 Accuracy: 0.9161 F1_score: 0.9159
         Validation - Loss: 0.2900 Accuracy: 0.8942 F1_score: 0.8940
Epoch: 23
         Train      - Loss: 0.2272 Accuracy: 0.9206 F1_score: 0.9205
         Validation - Loss: 0.2873 Accuracy: 0.8958 F1_score: 0.8958
Epoch: 24
         Train      - Loss: 0.2226 Accuracy: 0.9211 F1_score: 0.9209
         Validation - Loss: 0.2806 Accuracy: 0.8980 F1_score: 0.8981
Epoch: 25
         Train      - Loss: 0.2165 Accuracy: 0.9226 F1_score: 0.9225
         Validation - Loss: 0.2850 Accuracy: 0.8957 F1_score: 0.8956
Epoch: 26
         Train      - Loss: 0.2101 Accuracy: 0.9250 F1_score: 0.9249
         Validation - Loss: 0.2838 Accuracy: 0.8975 F1_score: 0.8975
Epoch: 27
         Train      - Loss: 0.2037 Accuracy: 0.9276 F1_score: 0.9275
         Validation - Loss: 0.2828 Accuracy: 0.8997 F1_score: 0.8993
Epoch: 28
         Train      - Loss: 0.2013 Accuracy: 0.9289 F1_score: 0.9288
         Validation - Loss: 0.2807 Accuracy: 0.8980 F1_score: 0.8979
Epoch: 29
         Train      - Loss: 0.1958 Accuracy: 0.9311 F1_score: 0.9310
         Validation - Loss: 0.2854 Accuracy: 0.8972 F1_score: 0.8971
Epoch: 30
         Train      - Loss: 0.1901 Accuracy: 0.9330 F1_score: 0.9329
         Validation - Loss: 0.2842 Accuracy: 0.8995 F1_score: 0.8997
<Figure size 432x288 with 0 Axes>
```




```
Loading model at epoch 30 for best validation f1
Preparing test loaders
Result on labelled test set  : Loss: 0.3700 Accuracy: 0.8691 F1_score: 0.8723
Result on unlabelled test set: Loss: 0.5518 Accuracy: 0.8797 F1_score: 0.8892
Result on full test set      : Loss: 0.4247 Accuracy: 0.8723 F1_score: 0.8733
```

Approach #3, use KMeans with Auto Encoder to compute label for unlabelled data

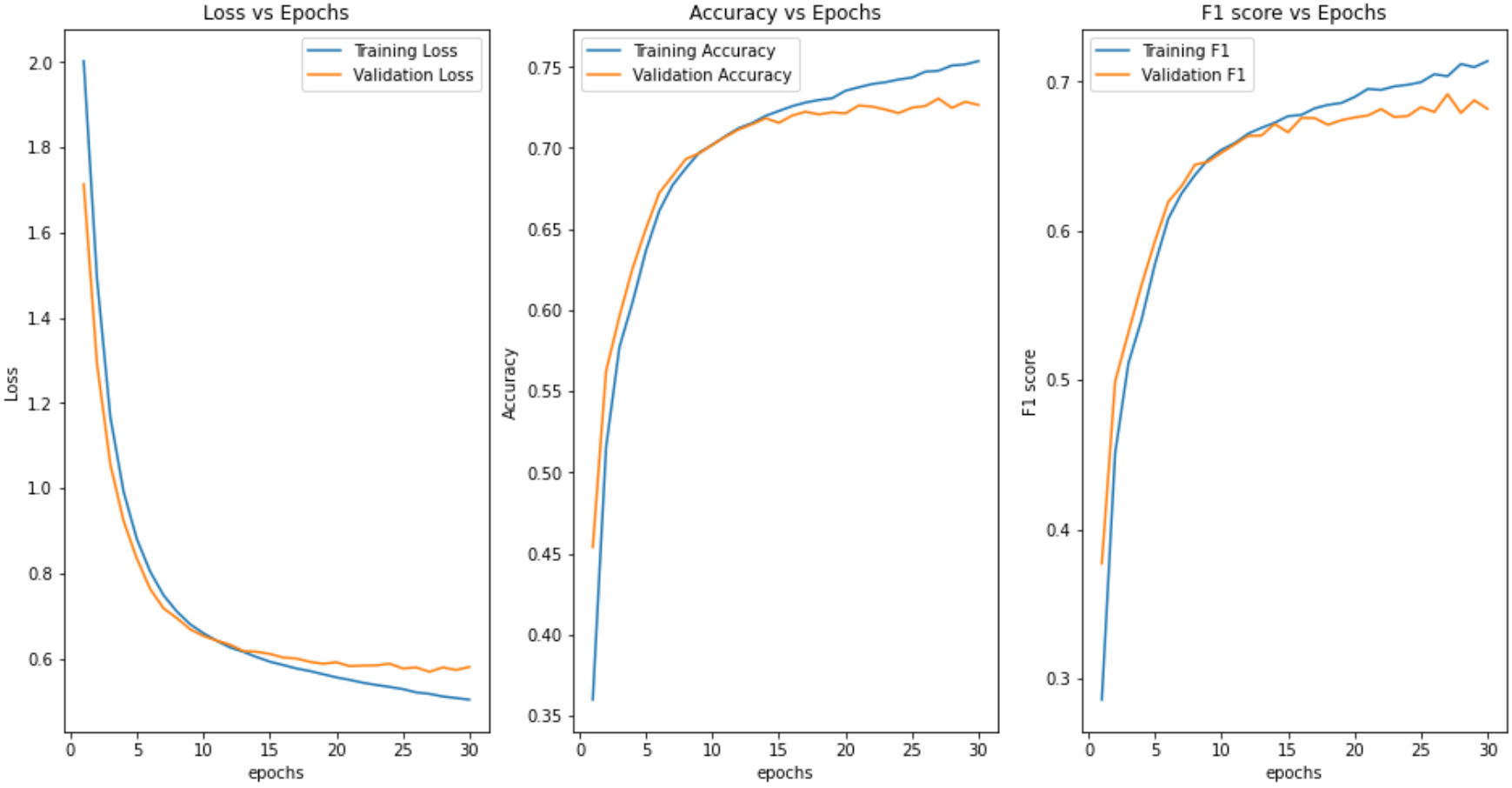
In [13]:
```
whole_flow(2, True)
```

```
Getting train and validate dataloaders for mode 2: clustering: kmeans with Auto Encoder
Labeling unlabeled data...
Training Auto Encoder...
/ssd/jason_ssd/AdaptToNewClass/mylibs/loss.py:8: UserWarning: To copy construct from a tensor, it is recommende
d to use sourceTensor.clone().detach() or sourceTensor.clone().detach().requires_grad_(True), rather than torc
```

```
h.tensor(sourceTensor).
  log2pi = torch.log(torch.tensor(2.0 * torch.as_tensor(np.pi)))
/ssd/jason_ssd/AdaptToNewClass/mylibs/loss.py:9: UserWarning: To copy construct from a tensor, it is recommende
d to use sourceTensor.clone().detach() or sourceTensor.clone().detach().requires_grad_(True), rather than torc
h.tensor(sourceTensor).
  return torch.sum(-0.5 * ((sample - mean) ** 2.0 * torch.exp(torch.tensor(-logvar)) + logvar + log2pi), dim=ra
xis)
Epoch: 1 Loss: {'loss': 356.0547530924479}
Epoch: 2 Loss: {'loss': 321.3297322265625}
Epoch: 3 Loss: {'loss': 318.8270221191406}
Epoch: 4 Loss: {'loss': 317.85027114257815}
Epoch: 5 Loss: {'loss': 317.0417849121094}
Epoch: 6 Loss: {'loss': 316.6882287109375}
Epoch: 7 Loss: {'loss': 316.0896029622396}
Epoch: 8 Loss: {'loss': 315.8089850911458}
Epoch: 9 Loss: {'loss': 315.3270469075521}
Epoch: 10 Loss: {'loss': 315.0863637532552}
Epoch: 11 Loss: {'loss': 315.12866446940103}
Epoch: 12 Loss: {'loss': 314.92135056966146}
Epoch: 13 Loss: {'loss': 314.81359272460935}
Epoch: 14 Loss: {'loss': 314.7329821126302}
Epoch: 15 Loss: {'loss': 314.628393766276}
Labeling accuracy: 0.3397777777777778
Start Training... ResNet
Epoch: 1
        Train      - Loss: 2.0030 Accuracy: 0.3599 F1_score: 0.2859
        Validation - Loss: 1.7137 Accuracy: 0.4540 F1_score: 0.3771
Epoch: 2
        Train      - Loss: 1.4915 Accuracy: 0.5155 F1_score: 0.4510
        Validation - Loss: 1.2921 Accuracy: 0.5625 F1_score: 0.4990
Epoch: 3
        Train      - Loss: 1.1675 Accuracy: 0.5771 F1_score: 0.5113
        Validation - Loss: 1.0559 Accuracy: 0.5962 F1_score: 0.5319
Epoch: 4
        Train      - Loss: 0.9913 Accuracy: 0.6051 F1_score: 0.5408
        Validation - Loss: 0.9226 Accuracy: 0.6260 F1_score: 0.5641
Epoch: 5
        Train      - Loss: 0.8799 Accuracy: 0.6367 F1_score: 0.5778
        Validation - Loss: 0.8338 Accuracy: 0.6503 F1_score: 0.5932
Epoch: 6
        Train      - Loss: 0.8038 Accuracy: 0.6613 F1_score: 0.6081
        Validation - Loss: 0.7629 Accuracy: 0.6723 F1_score: 0.6193
Epoch: 7
        Train      - Loss: 0.7485 Accuracy: 0.6771 F1_score: 0.6250
        Validation - Loss: 0.7175 Accuracy: 0.6827 F1_score: 0.6299
Epoch: 8
        Train      - Loss: 0.7101 Accuracy: 0.6875 F1_score: 0.6373
        Validation - Loss: 0.6947 Accuracy: 0.6932 F1_score: 0.6442
Epoch: 9
```

```
         Train      - Loss: 0.6803 Accuracy: 0.6971 F1_score: 0.6477
         Validation - Loss: 0.6684 Accuracy: 0.6967 F1_score: 0.6462
Epoch: 10
         Train      - Loss: 0.6590 Accuracy: 0.7020 F1_score: 0.6542
         Validation - Loss: 0.6527 Accuracy: 0.7020 F1_score: 0.6522
Epoch: 11
         Train      - Loss: 0.6414 Accuracy: 0.7074 F1_score: 0.6587
         Validation - Loss: 0.6416 Accuracy: 0.7070 F1_score: 0.6578
Epoch: 12
         Train      - Loss: 0.6256 Accuracy: 0.7123 F1_score: 0.6650
         Validation - Loss: 0.6324 Accuracy: 0.7115 F1_score: 0.6635
Epoch: 13
         Train      - Loss: 0.6153 Accuracy: 0.7154 F1_score: 0.6689
         Validation - Loss: 0.6171 Accuracy: 0.7147 F1_score: 0.6638
Epoch: 14
         Train      - Loss: 0.6032 Accuracy: 0.7198 F1_score: 0.6724
         Validation - Loss: 0.6155 Accuracy: 0.7183 F1_score: 0.6715
Epoch: 15
         Train      - Loss: 0.5921 Accuracy: 0.7228 F1_score: 0.6767
         Validation - Loss: 0.6105 Accuracy: 0.7155 F1_score: 0.6659
Epoch: 16
         Train      - Loss: 0.5845 Accuracy: 0.7258 F1_score: 0.6776
         Validation - Loss: 0.6020 Accuracy: 0.7200 F1_score: 0.6755
Epoch: 17
         Train      - Loss: 0.5763 Accuracy: 0.7280 F1_score: 0.6820
         Validation - Loss: 0.5996 Accuracy: 0.7223 F1_score: 0.6754
Epoch: 18
         Train      - Loss: 0.5702 Accuracy: 0.7296 F1_score: 0.6843
         Validation - Loss: 0.5919 Accuracy: 0.7207 F1_score: 0.6709
Epoch: 19
         Train      - Loss: 0.5627 Accuracy: 0.7306 F1_score: 0.6855
         Validation - Loss: 0.5873 Accuracy: 0.7220 F1_score: 0.6740
Epoch: 20
         Train      - Loss: 0.5554 Accuracy: 0.7353 F1_score: 0.6895
         Validation - Loss: 0.5909 Accuracy: 0.7213 F1_score: 0.6758
Epoch: 21
         Train      - Loss: 0.5495 Accuracy: 0.7374 F1_score: 0.6950
         Validation - Loss: 0.5818 Accuracy: 0.7262 F1_score: 0.6772
Epoch: 22
         Train      - Loss: 0.5427 Accuracy: 0.7394 F1_score: 0.6943
         Validation - Loss: 0.5831 Accuracy: 0.7255 F1_score: 0.6815
Epoch: 23
         Train      - Loss: 0.5375 Accuracy: 0.7405 F1_score: 0.6967
         Validation - Loss: 0.5835 Accuracy: 0.7237 F1_score: 0.6761
Epoch: 24
         Train      - Loss: 0.5330 Accuracy: 0.7422 F1_score: 0.6977
         Validation - Loss: 0.5875 Accuracy: 0.7215 F1_score: 0.6768
Epoch: 25
         Train      - Loss: 0.5278 Accuracy: 0.7434 F1_score: 0.6995
```

```
        Validation - Loss: 0.5762 Accuracy: 0.7248 F1_score: 0.6827
Epoch: 26
        Train      - Loss: 0.5201 Accuracy: 0.7470 F1_score: 0.7048
        Validation - Loss: 0.5788 Accuracy: 0.7258 F1_score: 0.6795
Epoch: 27
        Train      - Loss: 0.5166 Accuracy: 0.7476 F1_score: 0.7035
        Validation - Loss: 0.5684 Accuracy: 0.7305 F1_score: 0.6915
Epoch: 28
        Train      - Loss: 0.5105 Accuracy: 0.7508 F1_score: 0.7117
        Validation - Loss: 0.5788 Accuracy: 0.7247 F1_score: 0.6788
Epoch: 29
        Train      - Loss: 0.5068 Accuracy: 0.7515 F1_score: 0.7095
        Validation - Loss: 0.5727 Accuracy: 0.7285 F1_score: 0.6874
Epoch: 30
        Train      - Loss: 0.5032 Accuracy: 0.7535 F1_score: 0.7136
        Validation - Loss: 0.5800 Accuracy: 0.7265 F1_score: 0.6816
<Figure size 432x288 with 0 Axes>
```




```
Loading model at epoch 27 for best validation f1
```

```
Preparing test loaders
Result on labelled test set  : Loss: 0.3664 Accuracy: 0.8707 F1_score: 0.8736
Result on unlabelled test set: Loss: 1.1532 Accuracy: 0.3393 F1_score: 0.2093
Result on full test set      : Loss: 0.6043 Accuracy: 0.7113 F1_score: 0.6698
```

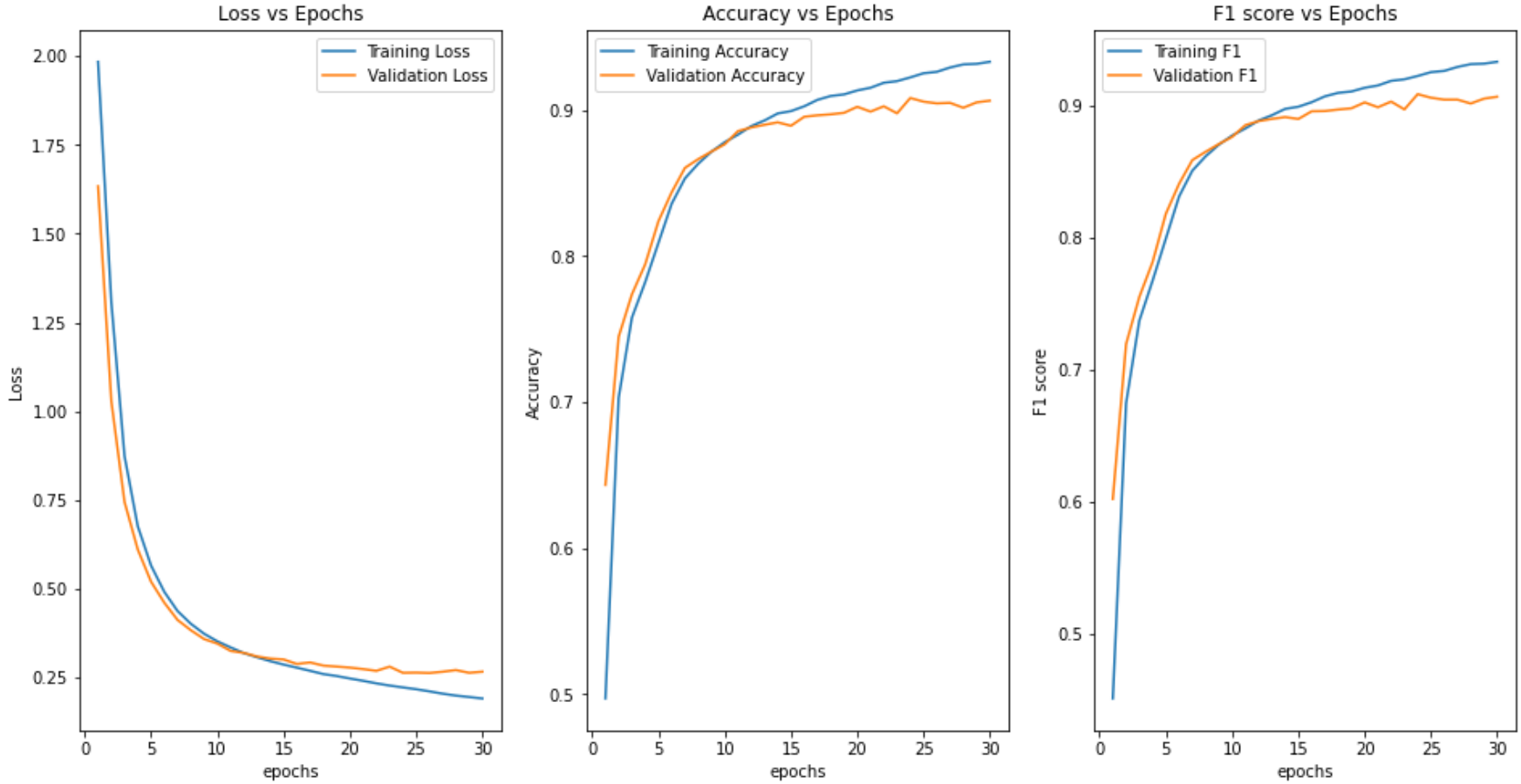
Approach #4, use Gaussian Mixture to compute label for unlabelled data

In [9]:
```
whole_flow(3, True)
```

```
Getting train and validate dataloaders for mode 3: clustering: Gaussian Mixture
Labeling unlabeled data...
Labeling accuracy: 0.8389444444444445
Start Training... ResNet
Epoch: 1
        Train      - Loss: 1.9841 Accuracy: 0.4972 F1_score: 0.4512
        Validation - Loss: 1.6337 Accuracy: 0.6433 F1_score: 0.6021
Epoch: 2
        Train      - Loss: 1.3083 Accuracy: 0.7031 F1_score: 0.6745
        Validation - Loss: 1.0294 Accuracy: 0.7448 F1_score: 0.7192
Epoch: 3
        Train      - Loss: 0.8740 Accuracy: 0.7579 F1_score: 0.7369
        Validation - Loss: 0.7450 Accuracy: 0.7738 F1_score: 0.7549
Epoch: 4
        Train      - Loss: 0.6764 Accuracy: 0.7821 F1_score: 0.7671
        Validation - Loss: 0.6110 Accuracy: 0.7943 F1_score: 0.7812
Epoch: 5
        Train      - Loss: 0.5661 Accuracy: 0.8090 F1_score: 0.7990
        Validation - Loss: 0.5205 Accuracy: 0.8238 F1_score: 0.8178
Epoch: 6
        Train      - Loss: 0.4920 Accuracy: 0.8358 F1_score: 0.8309
        Validation - Loss: 0.4614 Accuracy: 0.8437 F1_score: 0.8405
Epoch: 7
        Train      - Loss: 0.4379 Accuracy: 0.8531 F1_score: 0.8504
        Validation - Loss: 0.4129 Accuracy: 0.8603 F1_score: 0.8584
Epoch: 8
        Train      - Loss: 0.4021 Accuracy: 0.8632 F1_score: 0.8612
        Validation - Loss: 0.3840 Accuracy: 0.8663 F1_score: 0.8647
Epoch: 9
        Train      - Loss: 0.3741 Accuracy: 0.8714 F1_score: 0.8701
        Validation - Loss: 0.3591 Accuracy: 0.8715 F1_score: 0.8706
Epoch: 10
        Train      - Loss: 0.3527 Accuracy: 0.8779 F1_score: 0.8769
        Validation - Loss: 0.3464 Accuracy: 0.8765 F1_score: 0.8760
Epoch: 11
        Train      - Loss: 0.3355 Accuracy: 0.8831 F1_score: 0.8823
        Validation - Loss: 0.3257 Accuracy: 0.8855 F1_score: 0.8848
Epoch: 12
        Train      - Loss: 0.3196 Accuracy: 0.8889 F1_score: 0.8883
```

```
        Validation - Loss: 0.3201 Accuracy: 0.8880 F1_score: 0.8879
Epoch: 13
        Train      - Loss: 0.3077 Accuracy: 0.8928 F1_score: 0.8923
        Validation - Loss: 0.3099 Accuracy: 0.8898 F1_score: 0.8895
Epoch: 14
        Train      - Loss: 0.2969 Accuracy: 0.8976 F1_score: 0.8971
        Validation - Loss: 0.3041 Accuracy: 0.8915 F1_score: 0.8909
Epoch: 15
        Train      - Loss: 0.2873 Accuracy: 0.8992 F1_score: 0.8988
        Validation - Loss: 0.3016 Accuracy: 0.8892 F1_score: 0.8895
Epoch: 16
        Train      - Loss: 0.2785 Accuracy: 0.9025 F1_score: 0.9022
        Validation - Loss: 0.2890 Accuracy: 0.8953 F1_score: 0.8953
Epoch: 17
        Train      - Loss: 0.2697 Accuracy: 0.9069 F1_score: 0.9066
        Validation - Loss: 0.2932 Accuracy: 0.8963 F1_score: 0.8955
Epoch: 18
        Train      - Loss: 0.2605 Accuracy: 0.9096 F1_score: 0.9092
        Validation - Loss: 0.2844 Accuracy: 0.8970 F1_score: 0.8966
Epoch: 19
        Train      - Loss: 0.2550 Accuracy: 0.9106 F1_score: 0.9103
        Validation - Loss: 0.2816 Accuracy: 0.8980 F1_score: 0.8976
Epoch: 20
        Train      - Loss: 0.2480 Accuracy: 0.9133 F1_score: 0.9131
        Validation - Loss: 0.2786 Accuracy: 0.9022 F1_score: 0.9020
Epoch: 21
        Train      - Loss: 0.2415 Accuracy: 0.9152 F1_score: 0.9149
        Validation - Loss: 0.2745 Accuracy: 0.8988 F1_score: 0.8984
Epoch: 22
        Train      - Loss: 0.2344 Accuracy: 0.9186 F1_score: 0.9183
        Validation - Loss: 0.2696 Accuracy: 0.9025 F1_score: 0.9026
Epoch: 23
        Train      - Loss: 0.2282 Accuracy: 0.9197 F1_score: 0.9195
        Validation - Loss: 0.2814 Accuracy: 0.8977 F1_score: 0.8967
Epoch: 24
        Train      - Loss: 0.2229 Accuracy: 0.9223 F1_score: 0.9220
        Validation - Loss: 0.2640 Accuracy: 0.9082 F1_score: 0.9083
Epoch: 25
        Train      - Loss: 0.2178 Accuracy: 0.9251 F1_score: 0.9249
        Validation - Loss: 0.2646 Accuracy: 0.9057 F1_score: 0.9056
Epoch: 26
        Train      - Loss: 0.2118 Accuracy: 0.9261 F1_score: 0.9258
        Validation - Loss: 0.2635 Accuracy: 0.9045 F1_score: 0.9041
Epoch: 27
        Train      - Loss: 0.2053 Accuracy: 0.9290 F1_score: 0.9288
        Validation - Loss: 0.2671 Accuracy: 0.9048 F1_score: 0.9041
Epoch: 28
        Train      - Loss: 0.2000 Accuracy: 0.9312 F1_score: 0.9310
        Validation - Loss: 0.2717 Accuracy: 0.9015 F1_score: 0.9012
```

```
Epoch: 29
        Train      - Loss: 0.1958 Accuracy: 0.9315 F1_score: 0.9313
        Validation - Loss: 0.2641 Accuracy: 0.9052 F1_score: 0.9049
Epoch: 30
        Train      - Loss: 0.1918 Accuracy: 0.9330 F1_score: 0.9328
        Validation - Loss: 0.2672 Accuracy: 0.9063 F1_score: 0.9063
<Figure size 432x288 with 0 Axes>
```



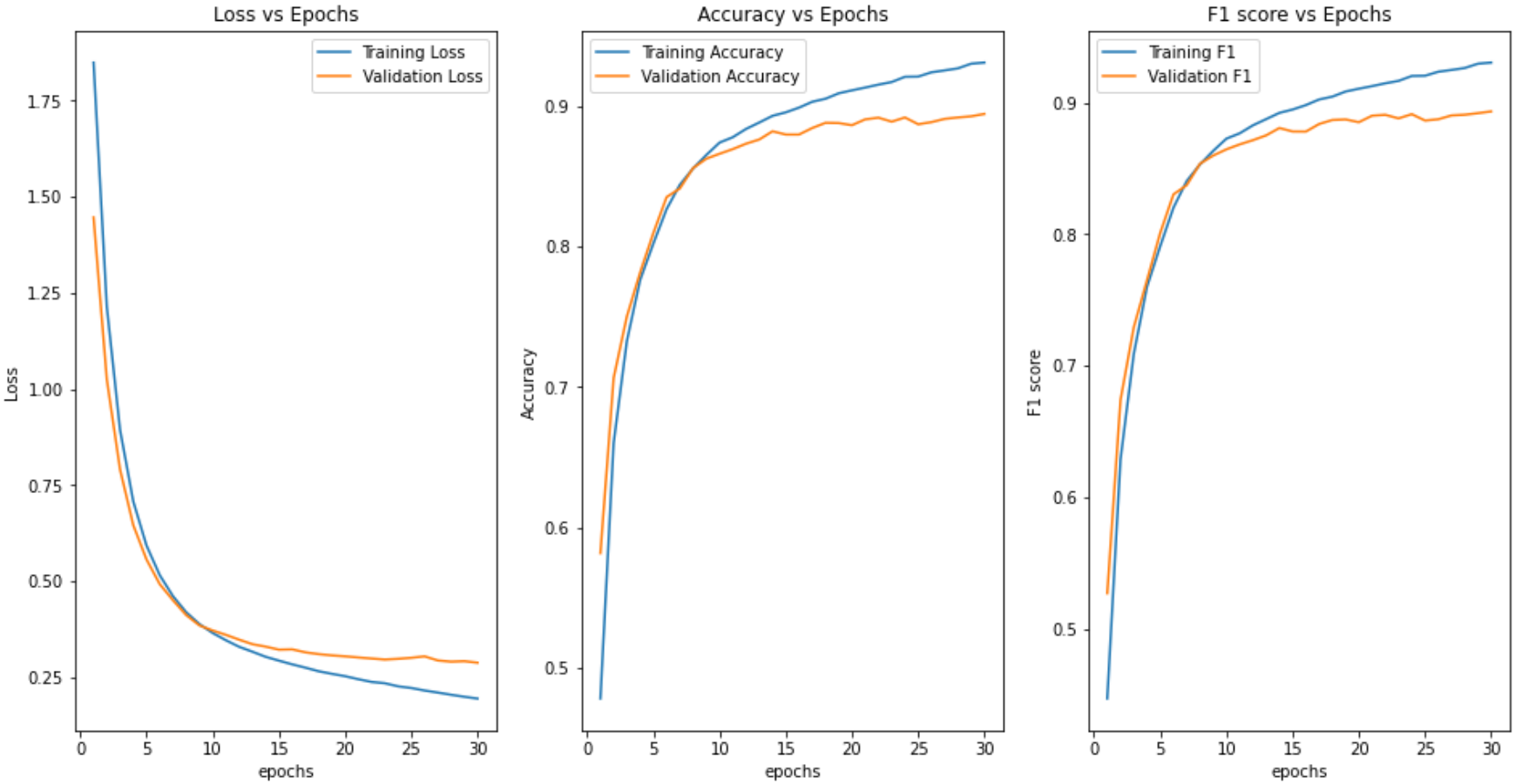

```
Loading model at epoch 24 for best validation f1
Preparing test loaders
Result on labelled test set  : Loss: 0.3543 Accuracy: 0.8730 F1_score: 0.8753
Result on unlabelled test set: Loss: 1.0503 Accuracy: 0.8207 F1_score: 0.8262
Result on full test set      : Loss: 0.5620 Accuracy: 0.8573 F1_score: 0.8565
```

Approach #5, use Gaussian Mixture with PCA to compute label for unlabelled data

In [10]:
```
whole_flow(4, True)
```

```
Getting train and validate dataloaders for mode 4: clustering: Gaussian Mixture with PCA
Labeling unlabeled data...
Labeling accuracy: 0.6936111111111111
Start Training... ResNet
Epoch: 1
        Train      - Loss: 1.8489 Accuracy: 0.4782 F1_score: 0.4468
        Validation - Loss: 1.4464 Accuracy: 0.5817 F1_score: 0.5269
Epoch: 2
        Train      - Loss: 1.2142 Accuracy: 0.6601 F1_score: 0.6288
        Validation - Loss: 1.0253 Accuracy: 0.7063 F1_score: 0.6739
Epoch: 3
        Train      - Loss: 0.8941 Accuracy: 0.7327 F1_score: 0.7087
        Validation - Loss: 0.7910 Accuracy: 0.7502 F1_score: 0.7291
Epoch: 4
        Train      - Loss: 0.7079 Accuracy: 0.7761 F1_score: 0.7600
        Validation - Loss: 0.6461 Accuracy: 0.7813 F1_score: 0.7649
Epoch: 5
        Train      - Loss: 0.5930 Accuracy: 0.8020 F1_score: 0.7909
        Validation - Loss: 0.5568 Accuracy: 0.8095 F1_score: 0.8010
Epoch: 6
        Train      - Loss: 0.5160 Accuracy: 0.8265 F1_score: 0.8202
        Validation - Loss: 0.4926 Accuracy: 0.8350 F1_score: 0.8303
Epoch: 7
        Train      - Loss: 0.4613 Accuracy: 0.8438 F1_score: 0.8405
        Validation - Loss: 0.4508 Accuracy: 0.8412 F1_score: 0.8373
Epoch: 8
        Train      - Loss: 0.4195 Accuracy: 0.8554 F1_score: 0.8531
        Validation - Loss: 0.4122 Accuracy: 0.8557 F1_score: 0.8536
Epoch: 9
        Train      - Loss: 0.3888 Accuracy: 0.8650 F1_score: 0.8634
        Validation - Loss: 0.3853 Accuracy: 0.8623 F1_score: 0.8599
Epoch: 10
        Train      - Loss: 0.3656 Accuracy: 0.8738 F1_score: 0.8727
        Validation - Loss: 0.3722 Accuracy: 0.8658 F1_score: 0.8646
Epoch: 11
        Train      - Loss: 0.3468 Accuracy: 0.8776 F1_score: 0.8768
        Validation - Loss: 0.3608 Accuracy: 0.8692 F1_score: 0.8683
Epoch: 12
        Train      - Loss: 0.3297 Accuracy: 0.8835 F1_score: 0.8829
        Validation - Loss: 0.3483 Accuracy: 0.8730 F1_score: 0.8716
Epoch: 13
        Train      - Loss: 0.3170 Accuracy: 0.8881 F1_score: 0.8876
        Validation - Loss: 0.3366 Accuracy: 0.8760 F1_score: 0.8752
Epoch: 14
        Train      - Loss: 0.3038 Accuracy: 0.8928 F1_score: 0.8923
        Validation - Loss: 0.3303 Accuracy: 0.8818 F1_score: 0.8808
Epoch: 15
```

```
        Train      - Loss: 0.2938 Accuracy: 0.8953 F1_score: 0.8949
        Validation - Loss: 0.3224 Accuracy: 0.8795 F1_score: 0.8782
Epoch: 16
        Train      - Loss: 0.2843 Accuracy: 0.8987 F1_score: 0.8983
        Validation - Loss: 0.3232 Accuracy: 0.8795 F1_score: 0.8781
Epoch: 17
        Train      - Loss: 0.2755 Accuracy: 0.9028 F1_score: 0.9025
        Validation - Loss: 0.3155 Accuracy: 0.8842 F1_score: 0.8839
Epoch: 18
        Train      - Loss: 0.2663 Accuracy: 0.9050 F1_score: 0.9047
        Validation - Loss: 0.3108 Accuracy: 0.8878 F1_score: 0.8869
Epoch: 19
        Train      - Loss: 0.2596 Accuracy: 0.9089 F1_score: 0.9086
        Validation - Loss: 0.3075 Accuracy: 0.8877 F1_score: 0.8873
Epoch: 20
        Train      - Loss: 0.2534 Accuracy: 0.9110 F1_score: 0.9107
        Validation - Loss: 0.3051 Accuracy: 0.8862 F1_score: 0.8852
Epoch: 21
        Train      - Loss: 0.2455 Accuracy: 0.9129 F1_score: 0.9127
        Validation - Loss: 0.3021 Accuracy: 0.8903 F1_score: 0.8902
Epoch: 22
        Train      - Loss: 0.2385 Accuracy: 0.9150 F1_score: 0.9148
        Validation - Loss: 0.2994 Accuracy: 0.8915 F1_score: 0.8908
Epoch: 23
        Train      - Loss: 0.2351 Accuracy: 0.9169 F1_score: 0.9167
        Validation - Loss: 0.2967 Accuracy: 0.8887 F1_score: 0.8881
Epoch: 24
        Train      - Loss: 0.2271 Accuracy: 0.9206 F1_score: 0.9204
        Validation - Loss: 0.2986 Accuracy: 0.8917 F1_score: 0.8913
Epoch: 25
        Train      - Loss: 0.2227 Accuracy: 0.9208 F1_score: 0.9206
        Validation - Loss: 0.3012 Accuracy: 0.8868 F1_score: 0.8865
Epoch: 26
        Train      - Loss: 0.2162 Accuracy: 0.9238 F1_score: 0.9236
        Validation - Loss: 0.3049 Accuracy: 0.8883 F1_score: 0.8875
Epoch: 27
        Train      - Loss: 0.2109 Accuracy: 0.9251 F1_score: 0.9250
        Validation - Loss: 0.2943 Accuracy: 0.8907 F1_score: 0.8904
Epoch: 28
        Train      - Loss: 0.2052 Accuracy: 0.9266 F1_score: 0.9264
        Validation - Loss: 0.2912 Accuracy: 0.8917 F1_score: 0.8909
Epoch: 29
        Train      - Loss: 0.1999 Accuracy: 0.9299 F1_score: 0.9297
        Validation - Loss: 0.2925 Accuracy: 0.8925 F1_score: 0.8921
Epoch: 30
        Train      - Loss: 0.1952 Accuracy: 0.9307 F1_score: 0.9306
        Validation - Loss: 0.2885 Accuracy: 0.8942 F1_score: 0.8934
<Figure size 432x288 with 0 Axes>
```




```
Loading model at epoch 30 for best validation f1
Preparing test loaders
Result on labelled test set  : Loss: 0.3643 Accuracy: 0.8696 F1_score: 0.8714
Result on unlabelled test set: Loss: 2.0664 Accuracy: 0.6857 F1_score: 0.6730
Result on full test set      : Loss: 0.8737 Accuracy: 0.8144 F1_score: 0.8085
```

Approach #6, use Gaussian Mixture with Auto Encoder to compute label for unlabelled data

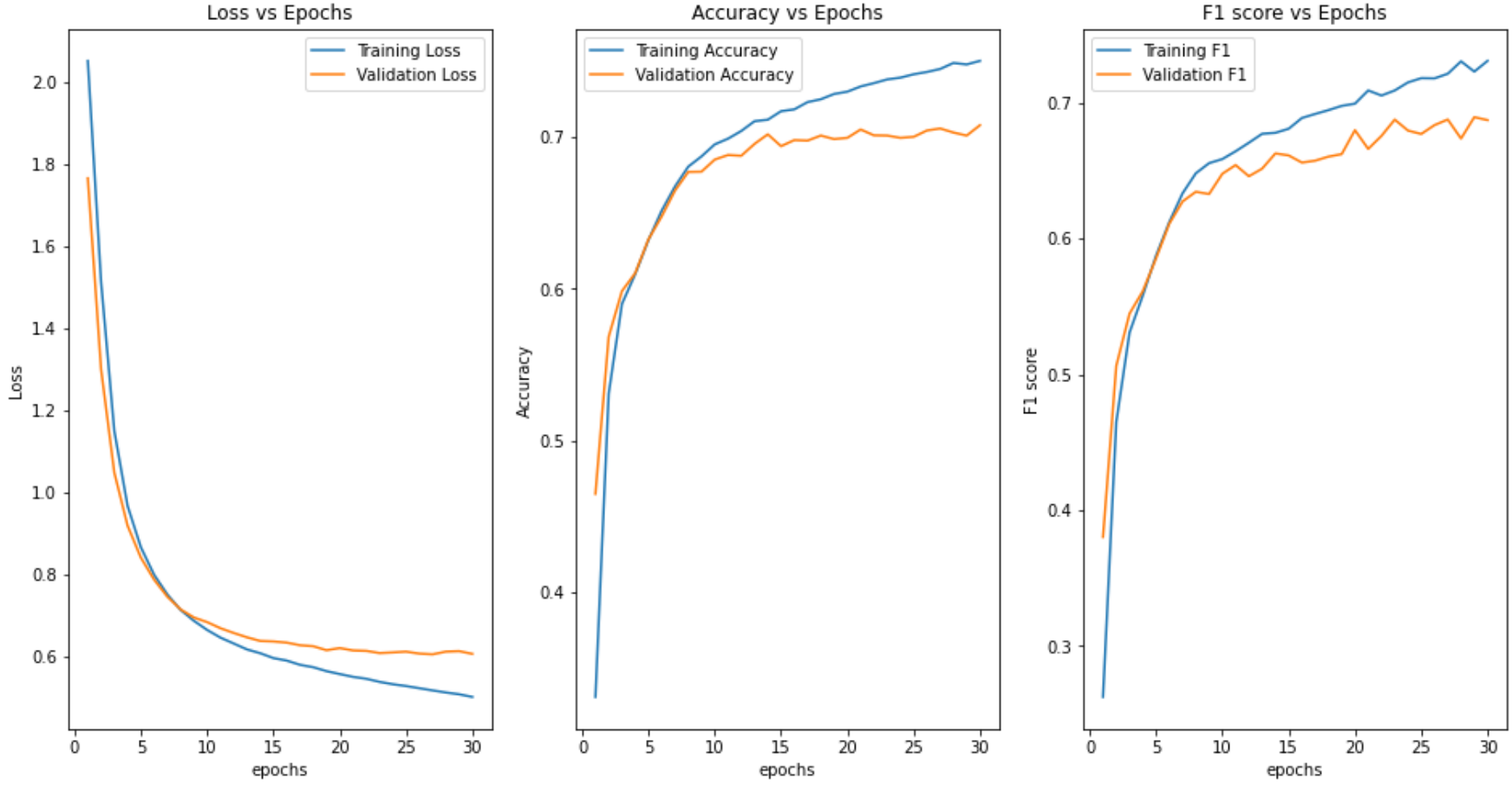
In [14]:
```
whole_flow(5, True)
```

```
Getting train and validate dataloaders for mode 5: clustering: Gaussian Mixture with Auto Encoder
Labeling unlabeled data...
Training Auto Encoder...
/ssd/jason_ssd/AdaptToNewClass/mylibs/loss.py:8: UserWarning: To copy construct from a tensor, it is recommende
d to use sourceTensor.clone().detach() or sourceTensor.clone().detach().requires_grad_(True), rather than torc
```

```
h.tensor(sourceTensor).
  log2pi = torch.log(torch.tensor(2.0 * torch.as_tensor(np.pi)))
/ssd/jason_ssd/AdaptToNewClass/mylibs/loss.py:9: UserWarning: To copy construct from a tensor, it is recommende
d to use sourceTensor.clone().detach() or sourceTensor.clone().detach().requires_grad_(True), rather than torc
h.tensor(sourceTensor).
  return torch.sum(-0.5 * ((sample - mean) ** 2.0 * torch.exp(torch.tensor(-logvar)) + logvar + log2pi), dim=ra
xis)
Epoch: 1 Loss: {'loss': 368.93738191731774}
Epoch: 2 Loss: {'loss': 320.82726127929686}
Epoch: 3 Loss: {'loss': 317.4619801269531}
Epoch: 4 Loss: {'loss': 316.2859282389323}
Epoch: 5 Loss: {'loss': 315.9476617675781}
Epoch: 6 Loss: {'loss': 315.54754375}
Epoch: 7 Loss: {'loss': 315.3139657714844}
Epoch: 8 Loss: {'loss': 315.0032361328125}
Epoch: 9 Loss: {'loss': 314.6484292317708}
Epoch: 10 Loss: {'loss': 314.68175266927085}
Epoch: 11 Loss: {'loss': 314.27091466471353}
Epoch: 12 Loss: {'loss': 314.34442755533854}
Epoch: 13 Loss: {'loss': 314.3051361002604}
Epoch: 14 Loss: {'loss': 314.1194011230469}
Epoch: 15 Loss: {'loss': 313.98986388346356}
Labeling accuracy: 0.3362222222222222
Start Training... ResNet
Epoch: 1
        Train      - Loss: 2.0507 Accuracy: 0.3306 F1_score: 0.2624
        Validation - Loss: 1.7648 Accuracy: 0.4647 F1_score: 0.3802
Epoch: 2
        Train      - Loss: 1.5125 Accuracy: 0.5307 F1_score: 0.4649
        Validation - Loss: 1.2995 Accuracy: 0.5683 F1_score: 0.5063
Epoch: 3
        Train      - Loss: 1.1489 Accuracy: 0.5900 F1_score: 0.5308
        Validation - Loss: 1.0474 Accuracy: 0.5985 F1_score: 0.5448
Epoch: 4
        Train      - Loss: 0.9664 Accuracy: 0.6097 F1_score: 0.5581
        Validation - Loss: 0.9182 Accuracy: 0.6103 F1_score: 0.5613
Epoch: 5
        Train      - Loss: 0.8658 Accuracy: 0.6322 F1_score: 0.5874
        Validation - Loss: 0.8404 Accuracy: 0.6327 F1_score: 0.5855
Epoch: 6
        Train      - Loss: 0.7986 Accuracy: 0.6516 F1_score: 0.6122
        Validation - Loss: 0.7868 Accuracy: 0.6478 F1_score: 0.6110
Epoch: 7
        Train      - Loss: 0.7507 Accuracy: 0.6675 F1_score: 0.6334
        Validation - Loss: 0.7452 Accuracy: 0.6648 F1_score: 0.6274
Epoch: 8
        Train      - Loss: 0.7131 Accuracy: 0.6806 F1_score: 0.6481
        Validation - Loss: 0.7143 Accuracy: 0.6772 F1_score: 0.6345
Epoch: 9
```

```
         Train      - Loss: 0.6865 Accuracy: 0.6874 F1_score: 0.6556
         Validation - Loss: 0.6946 Accuracy: 0.6773 F1_score: 0.6329
Epoch: 10
         Train      - Loss: 0.6643 Accuracy: 0.6954 F1_score: 0.6586
         Validation - Loss: 0.6833 Accuracy: 0.6853 F1_score: 0.6478
Epoch: 11
         Train      - Loss: 0.6456 Accuracy: 0.6991 F1_score: 0.6643
         Validation - Loss: 0.6686 Accuracy: 0.6883 F1_score: 0.6542
Epoch: 12
         Train      - Loss: 0.6313 Accuracy: 0.7042 F1_score: 0.6706
         Validation - Loss: 0.6569 Accuracy: 0.6878 F1_score: 0.6459
Epoch: 13
         Train      - Loss: 0.6169 Accuracy: 0.7106 F1_score: 0.6772
         Validation - Loss: 0.6464 Accuracy: 0.6957 F1_score: 0.6516
Epoch: 14
         Train      - Loss: 0.6076 Accuracy: 0.7116 F1_score: 0.6780
         Validation - Loss: 0.6375 Accuracy: 0.7020 F1_score: 0.6628
Epoch: 15
         Train      - Loss: 0.5956 Accuracy: 0.7173 F1_score: 0.6809
         Validation - Loss: 0.6363 Accuracy: 0.6942 F1_score: 0.6613
Epoch: 16
         Train      - Loss: 0.5897 Accuracy: 0.7184 F1_score: 0.6889
         Validation - Loss: 0.6334 Accuracy: 0.6982 F1_score: 0.6560
Epoch: 17
         Train      - Loss: 0.5794 Accuracy: 0.7233 F1_score: 0.6919
         Validation - Loss: 0.6271 Accuracy: 0.6978 F1_score: 0.6574
Epoch: 18
         Train      - Loss: 0.5735 Accuracy: 0.7251 F1_score: 0.6947
         Validation - Loss: 0.6246 Accuracy: 0.7012 F1_score: 0.6604
Epoch: 19
         Train      - Loss: 0.5639 Accuracy: 0.7286 F1_score: 0.6978
         Validation - Loss: 0.6150 Accuracy: 0.6988 F1_score: 0.6622
Epoch: 20
         Train      - Loss: 0.5569 Accuracy: 0.7301 F1_score: 0.6994
         Validation - Loss: 0.6201 Accuracy: 0.6997 F1_score: 0.6799
Epoch: 21
         Train      - Loss: 0.5500 Accuracy: 0.7336 F1_score: 0.7092
         Validation - Loss: 0.6144 Accuracy: 0.7052 F1_score: 0.6661
Epoch: 22
         Train      - Loss: 0.5453 Accuracy: 0.7358 F1_score: 0.7054
         Validation - Loss: 0.6133 Accuracy: 0.7013 F1_score: 0.6757
Epoch: 23
         Train      - Loss: 0.5378 Accuracy: 0.7382 F1_score: 0.7092
         Validation - Loss: 0.6080 Accuracy: 0.7012 F1_score: 0.6877
Epoch: 24
         Train      - Loss: 0.5321 Accuracy: 0.7394 F1_score: 0.7151
         Validation - Loss: 0.6096 Accuracy: 0.6997 F1_score: 0.6796
Epoch: 25
         Train      - Loss: 0.5277 Accuracy: 0.7416 F1_score: 0.7182
```

```
        Validation - Loss: 0.6114 Accuracy: 0.7003 F1_score: 0.6771
Epoch: 26
        Train      - Loss: 0.5223 Accuracy: 0.7431 F1_score: 0.7181
        Validation - Loss: 0.6067 Accuracy: 0.7045 F1_score: 0.6836
Epoch: 27
        Train      - Loss: 0.5170 Accuracy: 0.7451 F1_score: 0.7215
        Validation - Loss: 0.6049 Accuracy: 0.7058 F1_score: 0.6878
Epoch: 28
        Train      - Loss: 0.5119 Accuracy: 0.7491 F1_score: 0.7306
        Validation - Loss: 0.6115 Accuracy: 0.7032 F1_score: 0.6737
Epoch: 29
        Train      - Loss: 0.5076 Accuracy: 0.7481 F1_score: 0.7231
        Validation - Loss: 0.6127 Accuracy: 0.7012 F1_score: 0.6895
Epoch: 30
        Train      - Loss: 0.5011 Accuracy: 0.7505 F1_score: 0.7310
        Validation - Loss: 0.6060 Accuracy: 0.7080 F1_score: 0.6873
<Figure size 432x288 with 0 Axes>
```



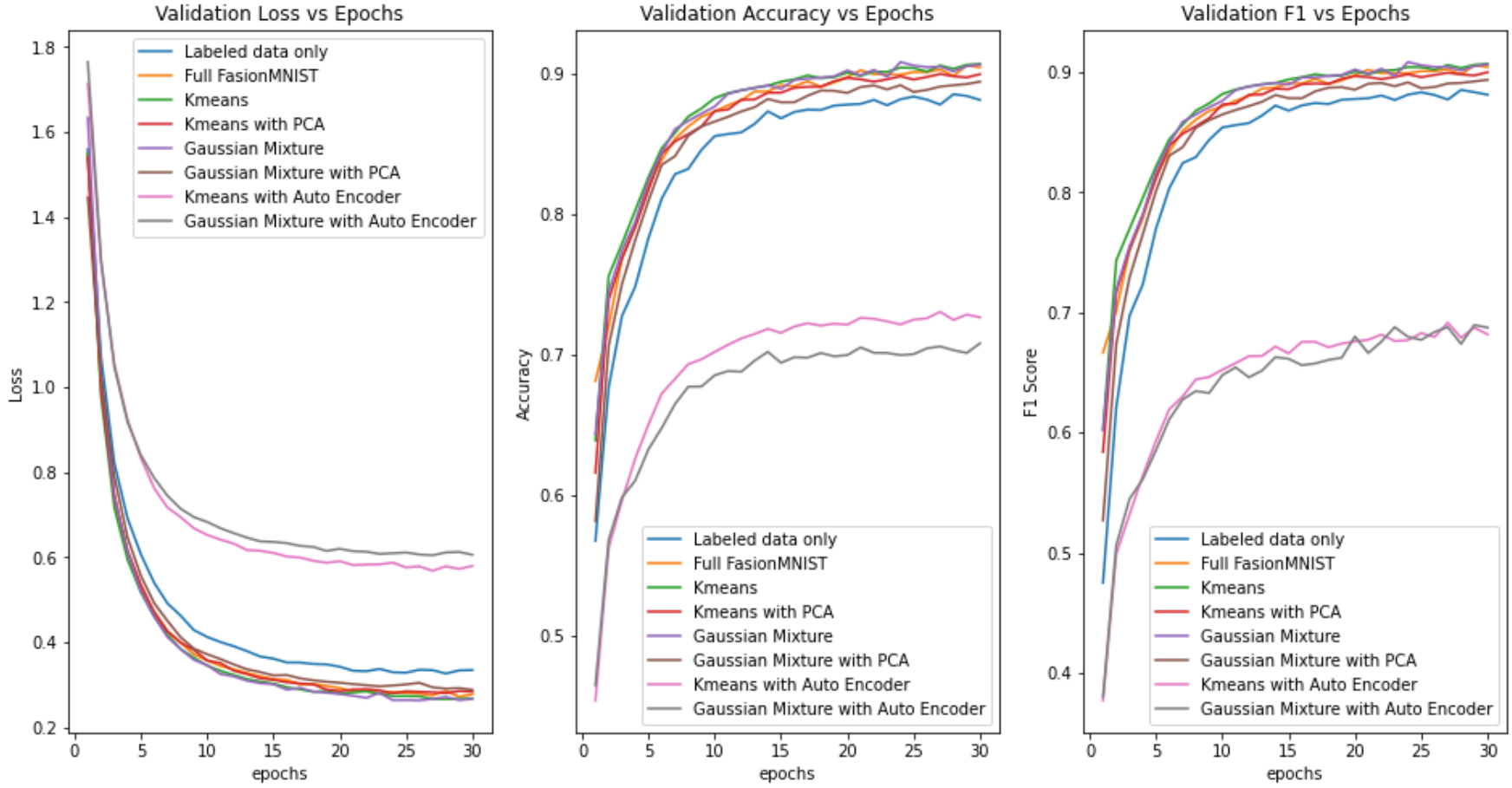

```
Loading model at epoch 29 for best validation f1
```

```
Preparing test loaders
Result on labelled test set  : Loss: 0.3663 Accuracy: 0.8691 F1_score: 0.8709
Result on unlabelled test set: Loss: 1.1304 Accuracy: 0.3223 F1_score: 0.2380
Result on full test set      : Loss: 0.5933 Accuracy: 0.7051 F1_score: 0.6783
```

Summary

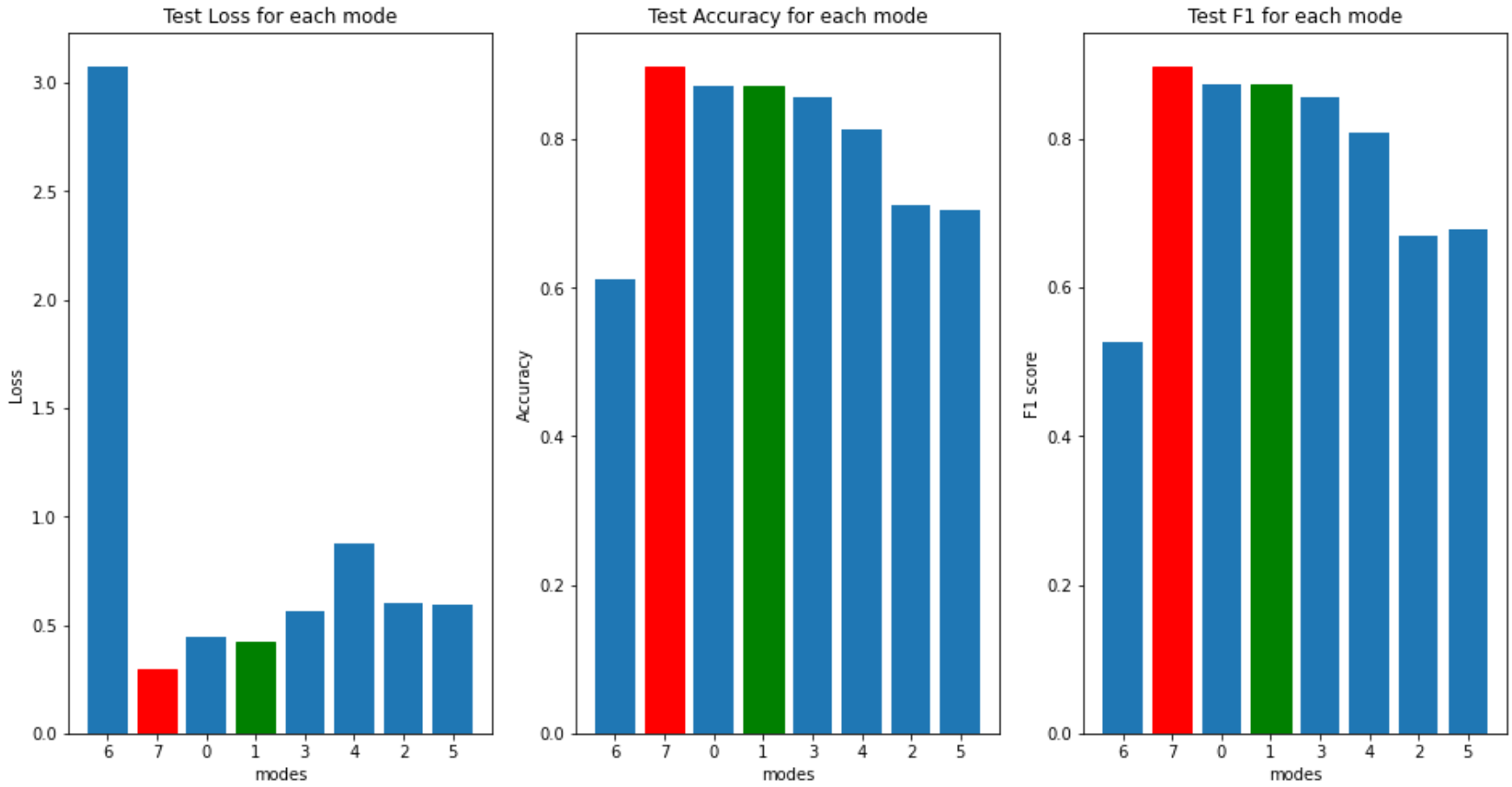
In [15]:
```
report_summary(mode_metrics, mode_description_short)
```

```
<Figure size 432x288 with 0 Axes>
```



In [18]:
```
report_test_summary(mode_test_metrics, mode_description_short)
```

```
<Figure size 432x288 with 0 Axes>
```




```
Mode Interpretations:
    6: Labeled data only
    7: Full FasionMNIST
    0: Kmeans
    1: Kmeans with PCA
    3: Gaussian Mixture
    4: Gaussian Mixture with PCA
    2: Kmeans with Auto Encoder
    5: Gaussian Mixture with Auto Encoder
Color Interpretations:
    Red  : Best result over all runs
    Green: Best result over all approachs (excluding baselines)
```

Method	Test Loss on Labeled Data	Test Accuracy on Labeled Data	Test F1 on Labeled Data	Test Loss on Unlabeled Data	Test Accuracy on Unlabeled Data	Test F1 on Unlabeled Data	Test Loss on All Data	Test Accuracy on All Data	Test F1 on All Data

Method	Test Loss on Labeled Data	Test Accuracy on Labeled Data	Test F1 on Labeled Data	Test Loss on Unlabeled Data	Test Accuracy on Unlabeled Data	Test F1 on Unlabeled Data	Test Loss on All Data	Test Accuracy on All Data	Test F1 on All Data
Baseline 1 - Only use Labeled data	0.3572	0.8726	0.8729	9.4141	0.0000	0.0000	3.0787	0.6108	0.5264
KMeans	0.3637	0.8716	0.8742	0.6130	0.8737	0.8823	0.4424	0.8722	0.8729
KMeans with PCA	0.3700	0.8691	0.8723	0.5518	0.8797	0.8892	0.4447	0.8723	0.8733
KMeans with VAE	0.3664	0.8707	0.8736	1.1532	0.3393	0.2093	0.6043	0.7113	0.6698
Gaussian Mixture	0.3543	0.8730	0.8753	1.0503	0.8207	0.8262	0.5620	0.8573	0.8565
Gaussian Mixture with PCA	0.3643	0.8696	0.8714	2.0664	0.6857	0.6730	0.8737	0.8144	0.8085
Gaussian Mixture with VAE	0.3663	0.8691	0.8709	1.1304	0.3223	0.2380	0.5933	0.7051	0.6783
Baseline 2 - Full FashionMNIST data	0.3777	0.8690	0.8719	0.0975	0.9683	0.9741	0.2934	0.8988	0.8984
Unsupervised Clustering	N/A	0.551	N/A	N/A	0.8876	N/A	N/A	0.5603	N/A

Conclusion

We implement two neural networks. One is with basic CNN, and the other is with pretrained ResNet18. The result for these two networks are similar. So we decide to draw conclusions based on the result from the basic CNN network.

We implement two clustering algorithms - KMeans and GaussianMixture, along with additional improvements - Principal Component Analysis and Variational Auto Encoder. Our implementation of VAE failed to properly encode the images. However, we can still see that without proper labeling, the overall model was still better than the one trained with only accurate labeled data and the unsupervised clustering. We can from here conclude that neural networks do not generalize to unseen data well or even at all, as we can see from the above table that when trained without certain labels, the model cannot recognize them at all (0 accuracy).

The model trained with only pre-labeled data has 60% accuracy among the full test set, which is a little better than the unsupervised clustering method with a 56% accuracy (the best is Kmean with PCA). The model trained with pre-labeled data is extremely bad on the unlabeled part of the test set. The accuracy of models trained with dataset combined by pre-labeled data and clustering labeled data are vary from 72% to 88%, which is getting close to the baseline2. Baseline2 is the best result with 91% accuracy. Fully labeled dataset is much accurate than the clustering labeled dataset. Thus, the best model is trained with the fully labeled one.

Among the methods we implemented, we see that KMeans generally performs slightly better than Gaussian Mixture. This is likely because of the properties of the dataset, where the clusters have clear decision boundaries. Principal Component Analysis can slightly improve the performance of the clustering labeling, but overall it doesn't affect the training result much.