Siamese_auxiliary

July 12, 2023

1 run load_data.ipynb BEFORE running this!

[11]:
```
import numpy as np
import torch
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
s = {
    'problem'           : "regression",
    'approach'          : "few-shot learning",
    'method'            : "non-parametric",
    'algorithm'         : "siamese network",
    'goal'              : "learn a distribution using few samples from it",
    'input'             : "samples from a distribution",
    'input type'        : "vectors",
    'input meaning'     : "spectrum",
    'output'            : "samples from a distribution",
    'output type'       : "one number",
    'output meaning'    : "temperature or pressure, depending on distribution",
    'number of ways'    : 2,
    'number of shots'   : 1,
    'number of folds'   : 8,
    'support-query ratio': 0.8,
    'task size'         : 5,
    'learning rate'     : 1e-4,
    'input dimension'   : 10000,
    'output dimension'  : 1,
    'feature dimension' : 300,
    'epoch'             : 1000,
    'epoch development' : 4,
    'data'              : 'pressure_230516_discrete',
    'cross validation round': 16,
    'cross validation round-development' : 3,
    'batch size'        : 32,
    'best model folder' : 'siamese_best_model/'
}
```

[12]:
```
import data_accessor as acc
data_names_list = [
```

```
    'temperature_230509_discrete',
    'pressure_230516_discrete'
    ]
data_dictionary = acc.setup(data_names_list)
```

```
loading temperature_230509_discrete____________________________
        input shape (number, dimension): (6000, 10000)
        label shape (number, dimension): (6000, 1)
        there are 16 folds
        4200 for training, 600 for validating, 1200 for testing
loading pressure_230516_discrete____________________________
        input shape (number, dimension): (5000, 10000)
        label shape (number, dimension): (5000, 1)
        there are 16 folds
        3500 for training, 500 for validating, 1000 for testing
```

[13]:
```
# task layout July 5, 2023

# siamese network extract feature space difference
# auxiliary network convert that difference into label difference

# Bing's feedback July 7th 2023
# need to fix:
# 1 convergence graph good for some CV round, bad for others
# 2 wrong use of cross-validation
# 3 try bigger dim, to fix 1
# 4 test on pressure data
# 5 try different architecture
#   if layer k has dimension d_k, d_k+1 should be > d_k / 10
```

[14]:
```
import torch.nn as nn
class SiameseNetwork(torch.nn.Module):
    def __init__(self, device, input_dimension, feature_dimension,␣
 ↪output_dimension):
        """ Input: input, anchor, anchor label
        Output: prediction for input"""
        super().__init__()
        self.input_dimension = input_dimension
        self.hidden_dimension = 600
        self.feature_hidden_dimension = 150
        self.feature_dimension = feature_dimension
        self.output_dimension = output_dimension
        self.device = device
        self.feature_sequential = torch.nn.Sequential(
            torch.nn.Linear(self.input_dimension, self.hidden_dimension),
            nn.ReLU(),
            torch.nn.Linear(self.hidden_dimension, self.hidden_dimension),
```

```
            nn.ReLU(),
            torch.nn.Linear(self.hidden_dimension, self.feature_dimension)
        )
        self.auxiliary_sequential = torch.nn.Sequential(
            torch.nn.Linear(self.feature_dimension, self.
↪feature_hidden_dimension),
            nn.ReLU(),
            torch.nn.Linear(self.feature_hidden_dimension, self.
↪feature_hidden_dimension),
            nn.ReLU(),
            torch.nn.Linear(self.feature_hidden_dimension, self.
↪output_dimension)
        )
        self.to(device)
        self.float()
    def forward(self, input, anchor, anchor_label):
        feature_input = self.feature_sequential(input)
        feature_anchor = self.feature_sequential(anchor)
        feature_space_difference_input_from_anchor = feature_input -␣
↪feature_anchor
        label_difference_input_from_anchor = self.
↪auxiliary_sequential(feature_space_difference_input_from_anchor)
        prediction = anchor_label + label_difference_input_from_anchor
        return prediction
```

[15]:
```
from tools import SaveBestModel, PatienceEarlyStopping, Scheduler, plot_loss
class Manager:
    """ DOES: train & evaluate a Siamese network
        """
    def __init__(self, epoch, cross_validation_round):
        self._network = SiameseNetwork(device, s['input dimension'], s['feature␣
↪dimension'], s['output dimension'])
        self._network.apply(self.initializer)
        self._learning_rate = s['learning rate']
        self._optimizer = torch.optim.Adam(
            params=self._network.parameters(), lr=self._learning_rate,
            weight_decay=3e-3)
        self._energy = nn.MSELoss()
        self._train_loss = []
        self._valid_loss = []
        self._test_loss = []
        self._epoch = epoch
        self._stopper = PatienceEarlyStopping(patience=5, min_delta=1e-7)
        self._cross_validation_round = cross_validation_round
        self._saver = SaveBestModel(s['best model folder'])
        self._scheduler = Scheduler(optimizer=self._optimizer,
```

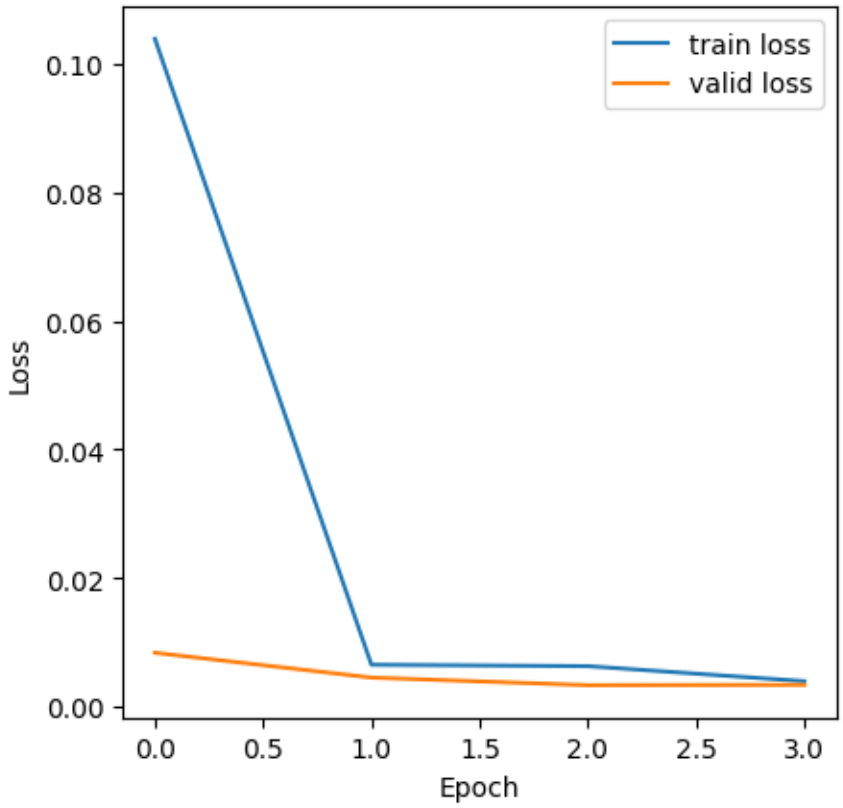
```
            minimum_learning_rate=1e-6, patience=5, factor=0.5)
    def initializer(self, layer):
        if type(layer) == nn.Linear:
            nn.init.kaiming_normal_(layer.weight) # normal version
    def _step(self, job):
        input, input_label, anchor, anchor_label = job
        # print(f"input dtype is {input_1.dtype}")
        prediction = self._network(input, anchor, anchor_label)
        loss = self._energy(input_label, prediction)
        return loss
    def train(self, train_dataloader, valid_dataloader):
        """ DOES: calculate loss from tasks
            NOTE: we have a BATCH of tasks here """
        for e in range(self._epoch):
            # print(f"train() epoch {e}")
            batch_train_loss = []
            for _, batch in enumerate(train_dataloader):
                self._optimizer.zero_grad()
                loss = self._step(batch)
                loss.backward()
                self._optimizer.step()
                batch_train_loss.append(loss.item())
            self._train_loss.append(np.mean(batch_train_loss))
            batch_valid_loss = []
            with torch.no_grad():
                for _, batch in enumerate(valid_dataloader):
                    loss = self._step(batch)
                    batch_valid_loss.append(loss.item())
            self._valid_loss.append(np.mean(batch_valid_loss))
            # saving, early stopping, scheduler for EACH epoch!
            self._saver(current_loss=np.mean(batch_valid_loss),
                  model=self._network,
                  round=self._cross_validation_round
                  )
            self._scheduler(np.mean(batch_valid_loss))
            self._stopper(np.mean(batch_valid_loss))
            if self._stopper.early_stop == True:
                print(f"EARLY STOPPING @ epoch {e}")
                break
        # summary printout, after we're done with epochs
        print(f"min train loss: {np.min(self._train_loss)}")
        print(f"min valid loss: {np.min(self._valid_loss)}")
        plot_loss(self._train_loss, self._valid_loss)
        return np.min(self._valid_loss)
    def test(self, test_dataloader):
        with torch.no_grad():
            batch_test_loss = []
```

```
            for _, batch in enumerate(test_dataloader):
                loss = self._step(batch)
                batch_test_loss.append(loss.item())
            self._test_loss.append(np.mean(batch_test_loss))
        return np.min(self._test_loss)
```

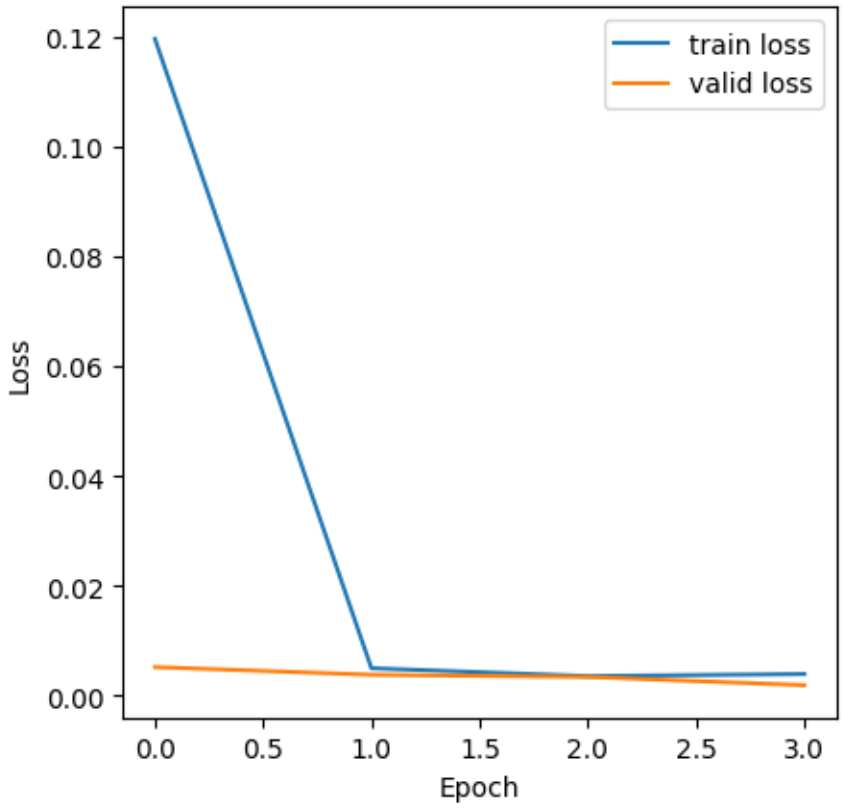
[16]:
```
from torch.utils.data import DataLoader
from tools import SiameseDataset, SaveBestCrossValidationModel

CV_saver = SaveBestCrossValidationModel(s['best model folder'])
test_indices = data_dictionary[s['data']]['test indices']
epoch = s['epoch development']
print(f"data: {s['data']}")
cross_validation_loss = []
for cross_validation_round, (train, valid) in enumerate(zip(
    data_dictionary[s['data']]['train indices'],
    data_dictionary[s['data']]['valid indices'])):
    if cross_validation_round < s['cross validation round']:
        print(f"CV round {cross_validation_round}")
        network_object = Manager(epoch, cross_validation_round)
        valid_loss = network_object.train(
            DataLoader(SiameseDataset(
            data_dictionary[s['data']]['data'],
            data_dictionary[s['data']]['label'],
            data_dictionary[s['data']]['train indices'][cross_validation_round],
            device=device,), shuffle=False, batch_size=s['batch size']),
            DataLoader(SiameseDataset(
            data_dictionary[s['data']]['data'],
            data_dictionary[s['data']]['label'],
            data_dictionary[s['data']]['valid indices'][cross_validation_round],
            device=device,), shuffle=False, batch_size=s['batch size']))
        CV_saver(current_loss=valid_loss, round=cross_validation_round)
        cross_validation_loss.append(valid_loss)
print()
print(f"\nbest model is: {CV_saver.best_model_name} with {CV_saver.
 ↪current_best_loss}")
print(f"The aggregate performance is: mean {np.mean(cross_validation_loss)},␣
 ↪std {np.std(cross_validation_loss)}")
```

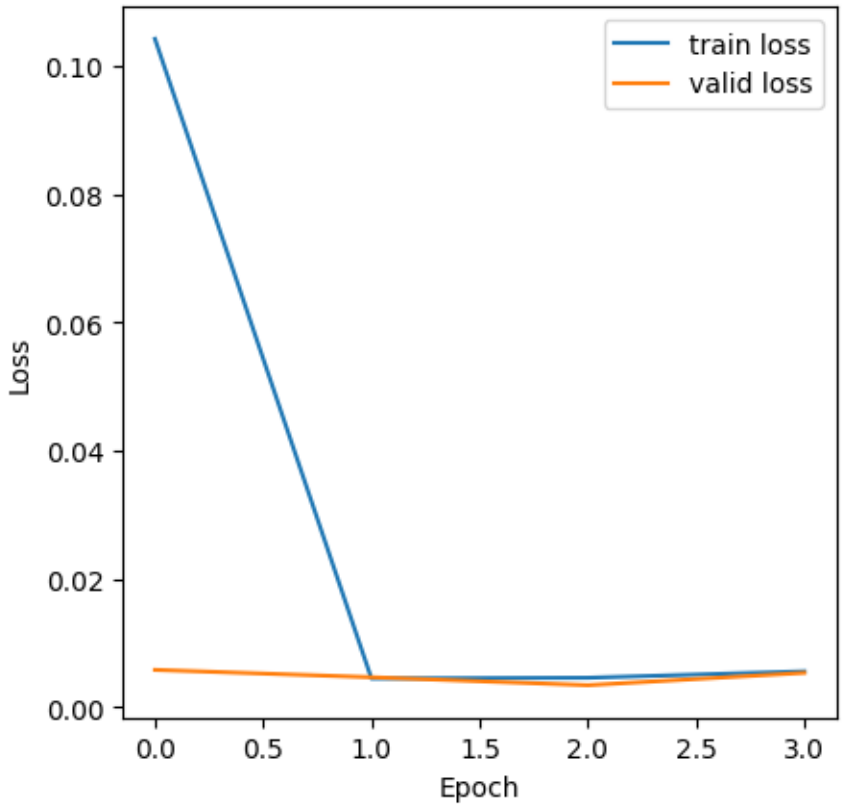
```
data: pressure_230516_discrete
CV round 0
min train loss: 0.0039208155569874425
min valid loss: 0.0032997974049067125
```



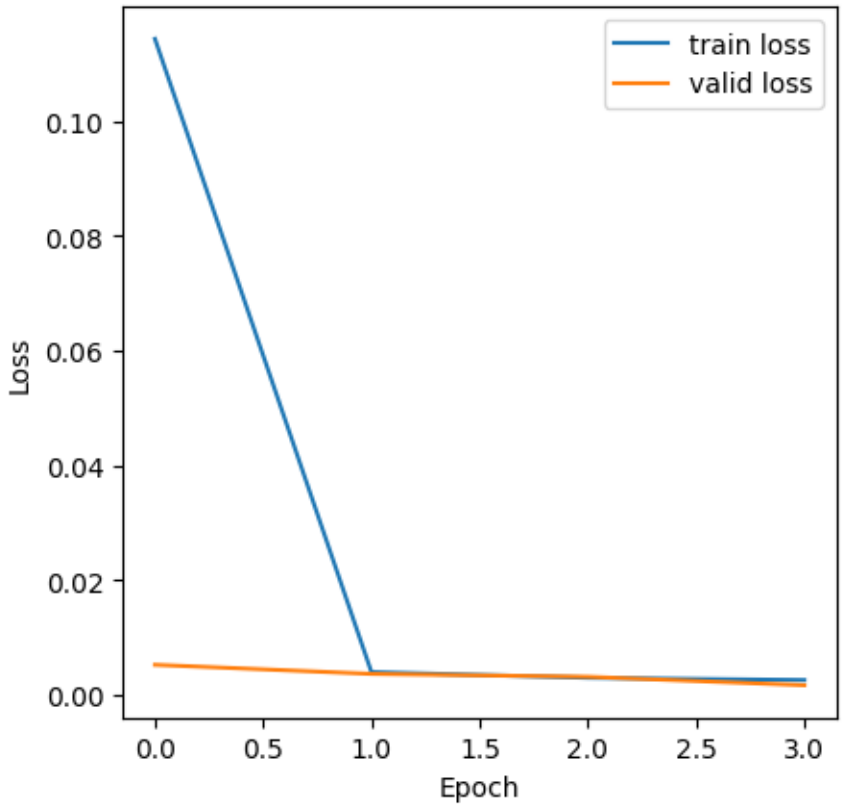

```
CV round 1
min train loss: 0.003589721716178412
min valid loss: 0.0019362719212949742
```



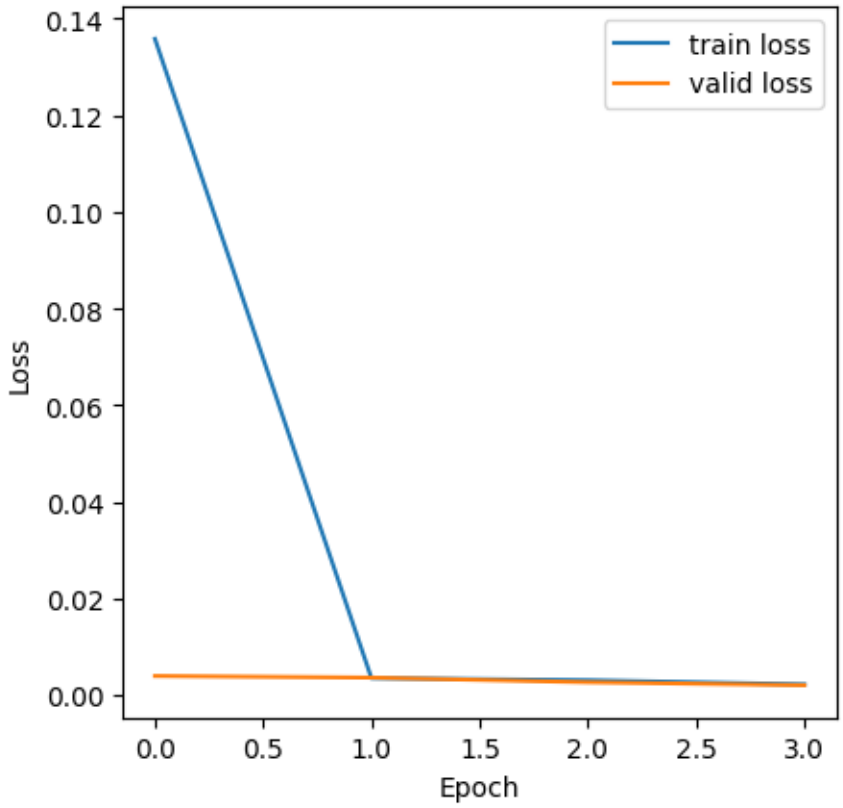

```
CV round 2
min train loss: 0.004447551108685068
min valid loss: 0.0033991358504863456
```



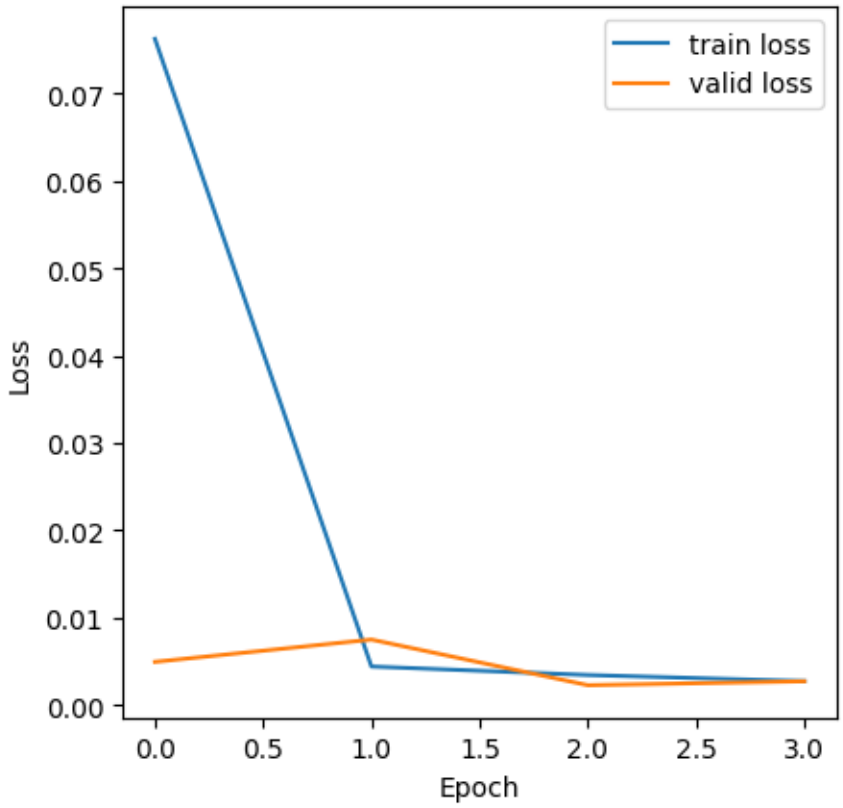

```
CV round 3
min train loss: 0.0025788469864478844
min valid loss: 0.0016827133658807725
```



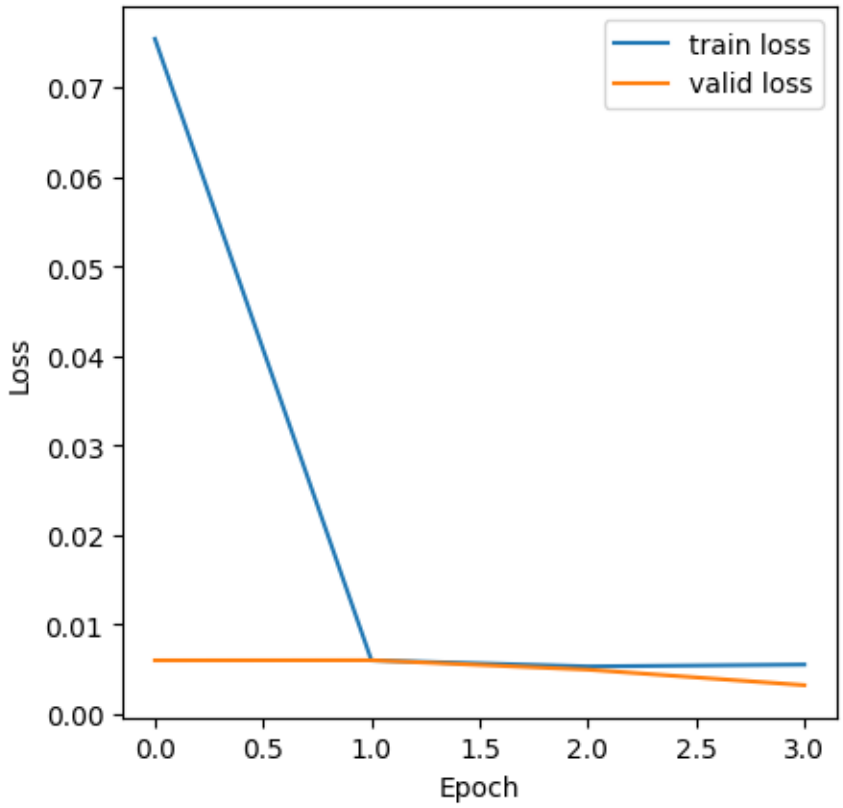

```
CV round 4
min train loss: 0.0022122193535324187
min valid loss: 0.002037715916230809
```



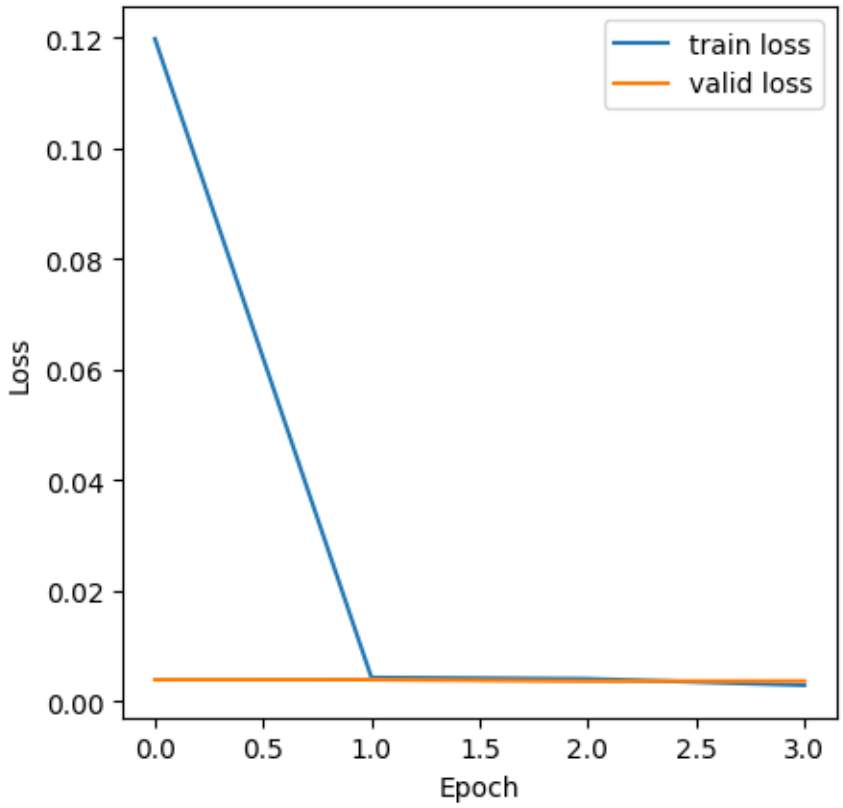

```
CV round 5
min train loss: 0.0027235894030044703
min valid loss: 0.002261809626361355
```



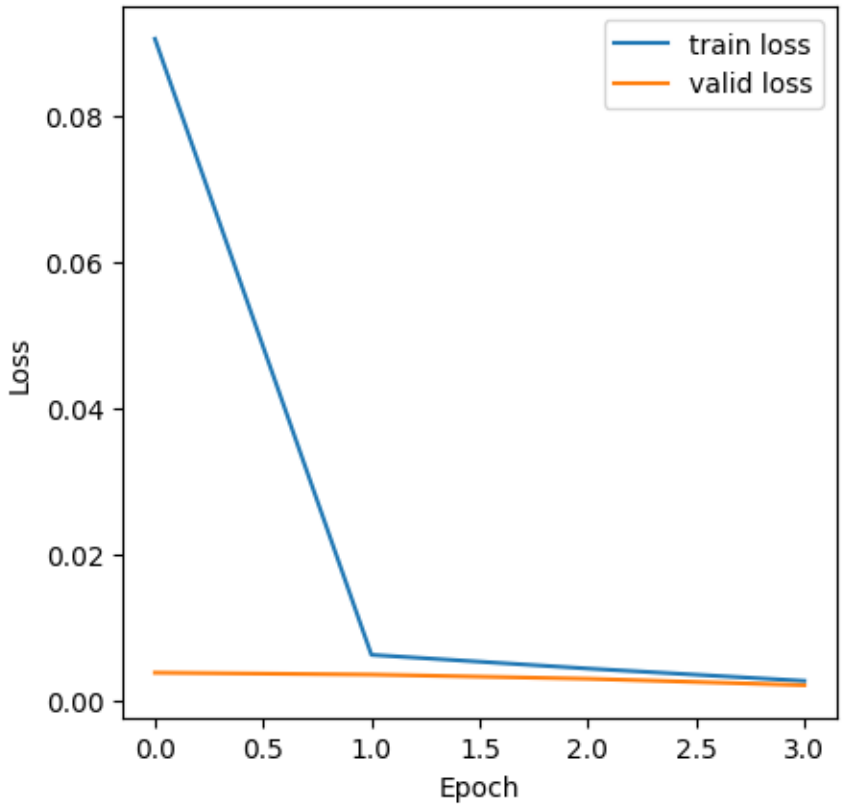

```
CV round 6
min train loss: 0.005532719307096506
min valid loss: 0.0032367689273087308
```



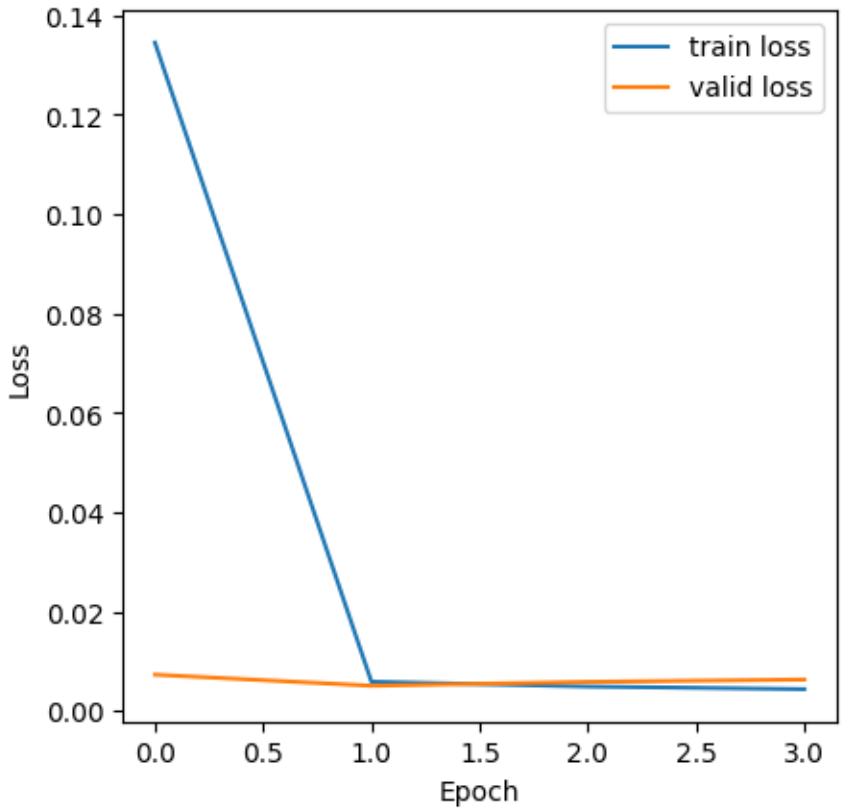

```
CV round 7
min train loss: 0.002919301141942428
min valid loss: 0.003652309562312439
```



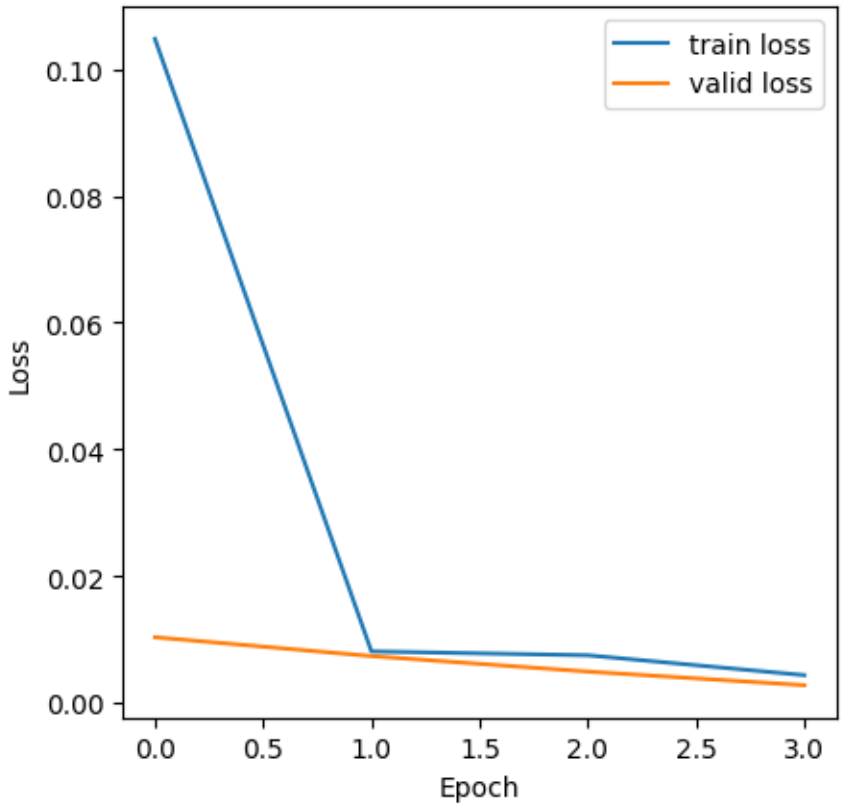

```
CV round 8
min train loss: 0.002585553759920665
min valid loss: 0.0019979974676971324
```



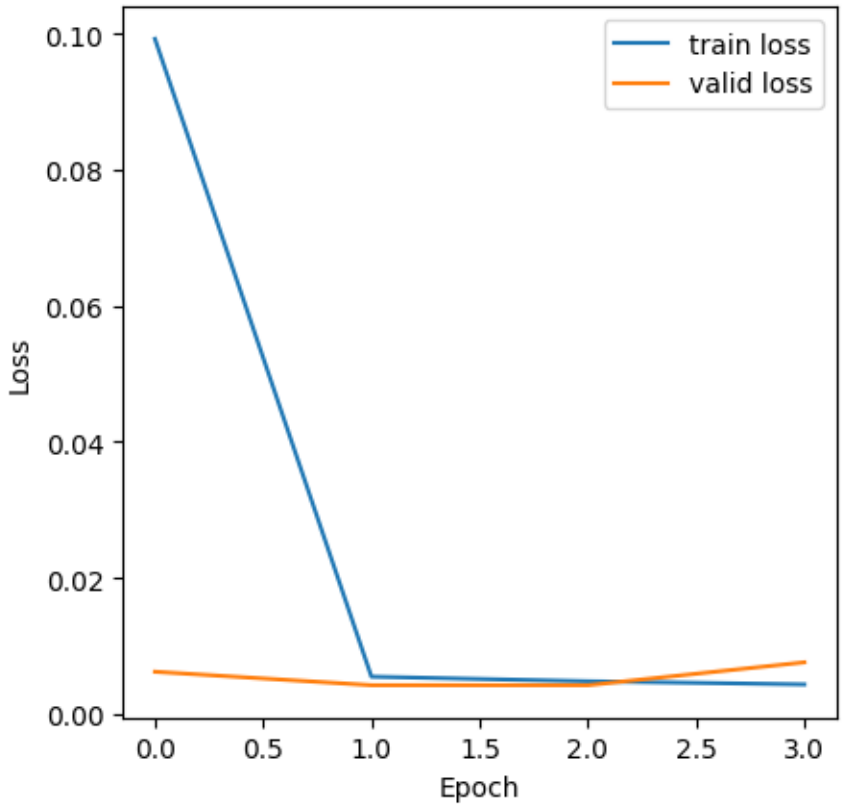

```
CV round 9
min train loss: 0.004387160411781886
min valid loss: 0.005081918410724029
```



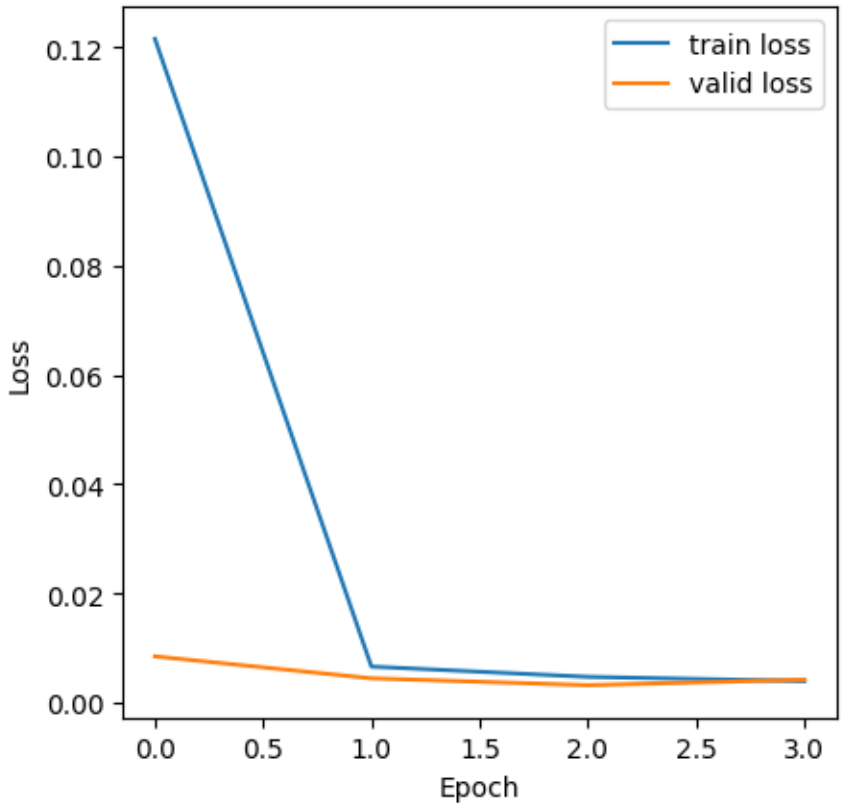

```
CV round 10
min train loss: 0.00435771146281199
min valid loss: 0.002774933527689427
```



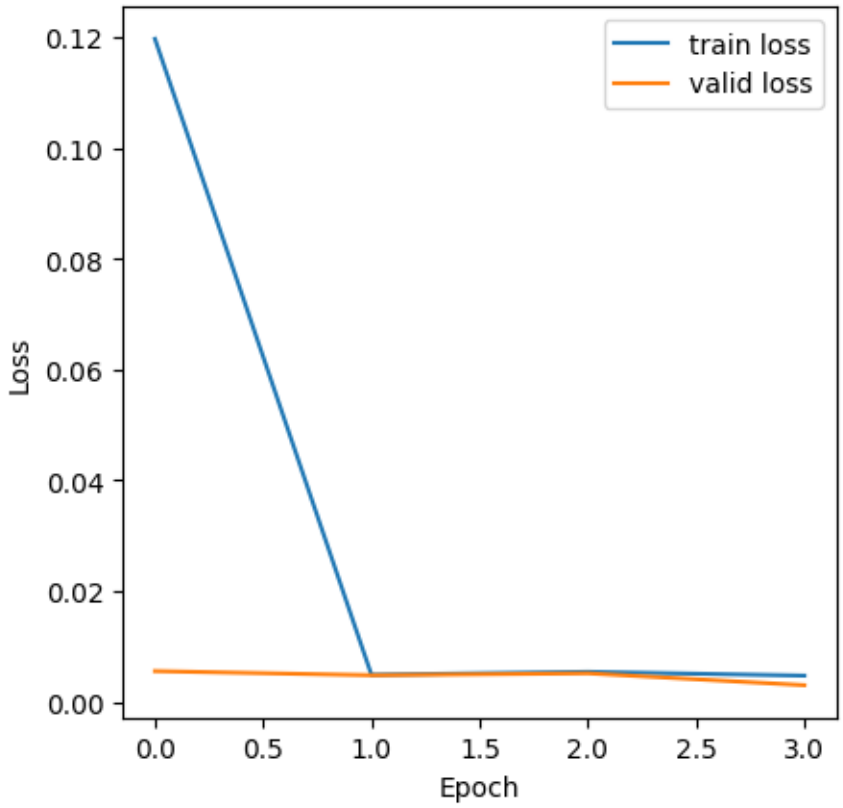

```
CV round 11
min train loss: 0.004356235781134191
min valid loss: 0.004240545182256028
```



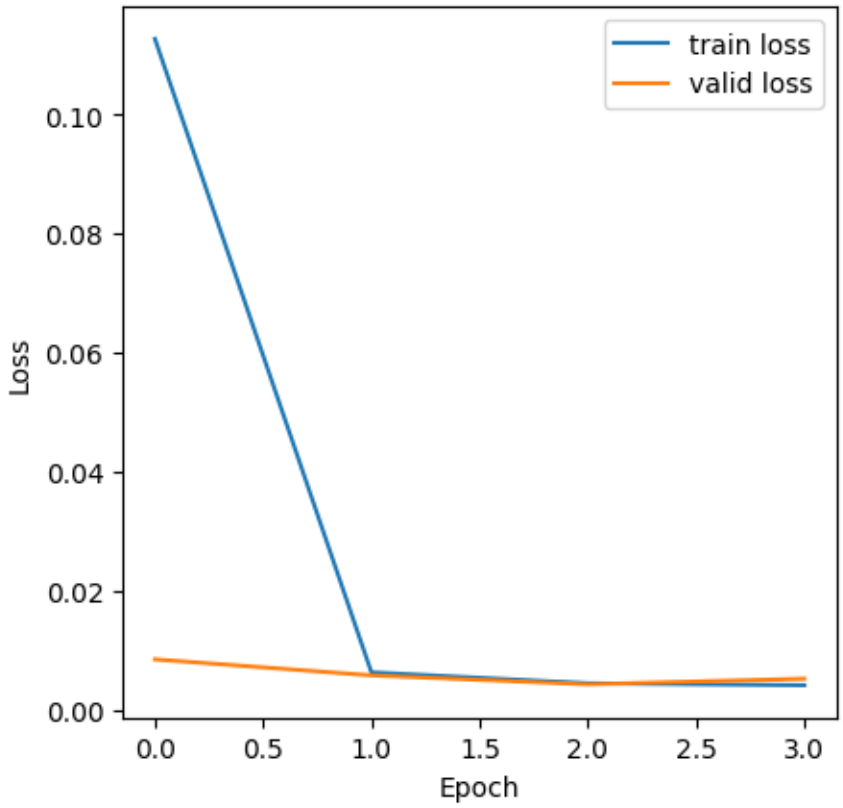

```
CV round 12
min train loss: 0.0039606870723549615
min valid loss: 0.0032080487144412473
```



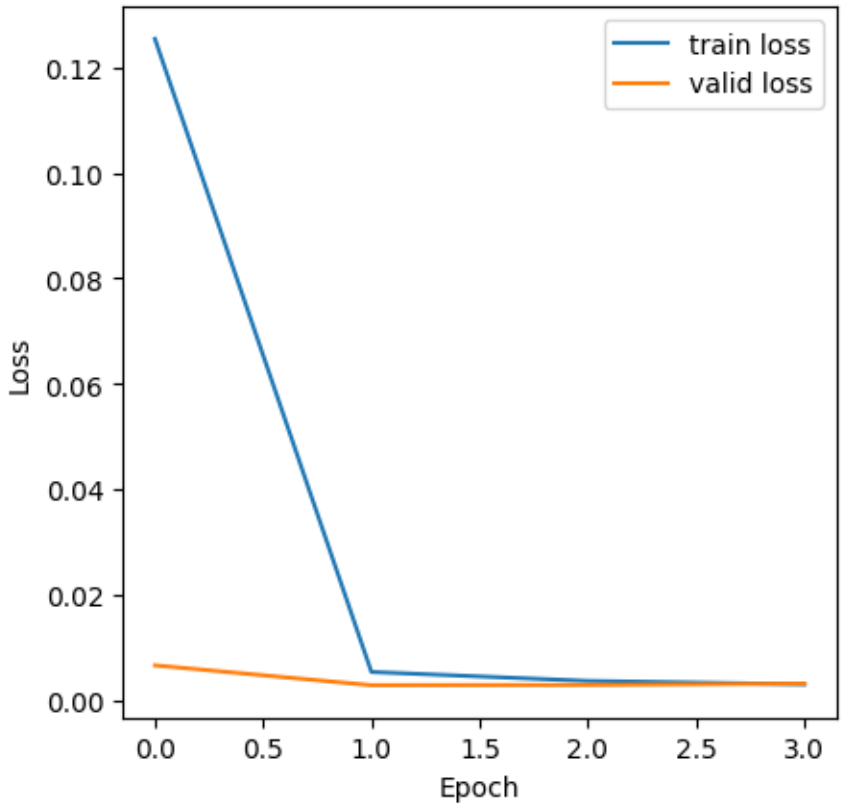

```
CV round 13
min train loss: 0.0047194683003577995
min valid loss: 0.0029927200521342456
```




```
CV round 14
min train loss: 0.004101748690432445
min valid loss: 0.004256487765815109
```




```
CV round 15
min train loss: 0.002965034574100917
min valid loss: 0.0027946264890488237
```




```
best model is: CV=3.pth with 0.0016827133658807725
The aggregate performance is: mean 0.0030533625115367613, std
0.0009224677882625907
```

[17]:
```
network_object._network.load_state_dict(torch.load(s['best model folder'] +␣
  ↪CV_saver.best_model_name))
test_loss = network_object.test(
            DataLoader(SiameseDataset(
            data_dictionary[s['data']]['data'],
            data_dictionary[s['data']]['label'],
            data_dictionary[s['data']]['test indices'],
            device=device,), shuffle=False, batch_size=s['batch size']))
print(f"testing loss: {test_loss}")
```

```
testing loss: 0.0018610085389809683
```