

Νευρωνικά Δίκτυα-Βαθιά Μάθηση

1^η Υποχρεωτική εργασία

Τεχνητή Αναφορά

Όνοματεπώνυμο: Ζαχαριάδης Αντώνιος

AEM: 4173

Στην παρούσα εργασία αναπτύχθηκε ένα νευρωνικό δίκτυο τύπου **MLP** (**Multilayer Perceptron**) με στόχο την ταξινόμηση εικόνων του συνόλου δεδομένων **CIFAR-10**, που περιλαμβάνει 10 κατηγορίες. Το δίκτυο εκπαιδεύτηκε με τη χρήση της συνάρτησης απώλειας **CrossEntropyLoss** και του βελτιστοποιητή **Adam**, αξιοποιώντας τεχνικές όπως η **κανονικοποίηση δεδομένων** και το **early stopping** για την αποφυγή υπερπροσαρμογής. Οι μετρικές ακρίβειας, ανάκλησης και F1-score χρησιμοποιήθηκαν για την αξιολόγηση της απόδοσης στο σύνολο δοκιμής. Η ανάλυση έδειξε βελτίωση της ακρίβειας με την προσαρμογή υπερπαραμέτρων, ενώ αναδείχθηκαν προκλήσεις στη γενίκευση του μοντέλου σε σύνθετα δεδομένα.

1) Περιγραφή κώδικα

```
# Μετασχηματισμοί για τη βάση δεδομένων (Κανονικοποίηση)
transform = transforms.Compose([
    transforms.ToTensor(),
    transforms.Normalize((0.5, 0.5, 0.5), (0.5, 0.5, 0.5))
])
```

Μετατρέπονται οι εικόνες σε tensors, δηλαδή σε πολυδιάστατους πίνακες κατάλληλους για το PyTorch και έπειτα κανονικοποιούνται. Το πρώτο όρισμα

της `transforms.Normalize(0.5, 0.5, 0.5)` είναι η μέση τιμή για κάθε κανάλι χρώματος (RGB). Το δεύτερο (0.5, 0.5, 0.5) είναι η τυπική απόκλιση.

```
# Φόρτωση CIFAR-10
train_dataset = torchvision.datasets.CIFAR10(root='./data', train=True, download=True, transform=transform)
test_dataset = torchvision.datasets.CIFAR10(root='./data', train=False, download=True, transform=transform)

batch_size = 90
train_loader = DataLoader(train_dataset, batch_size, shuffle=True) #Ανακάτεμα των δειγμάτων πριν δημιουργηθούν σε batches
test_loader = DataLoader(test_dataset, batch_size, shuffle=False)
```

Φορτώνεται το dataset της CIFAR-10.

- **`train=True / train=False`**: Επιλέγει αν φορτώνουμε τα δεδομένα εκπαίδευσης ή δοκιμής.
- **`download=True`**: Κατεβάζει αυτόματα τα δεδομένα αν δεν υπάρχουν ήδη
- **`shuffle=True/shuffle=False`**: Η παράμετρος `shuffle` στον ορισμό του `DataLoader` ελέγχει αν τα δεδομένα θα ανακατευτούν (`shuffled`) πριν δοθούν στο μοντέλο κατά την εκπαίδευση ή την αξιολόγηση πριν αυτά δημιουργηθούν σε batches.
- **`transform=transform`**: Εφαρμόζει τους μετασχηματισμούς που καθορίσαμε προηγουμένως.

```
class MLP(nn.Module): # Δήλωση της κλάσης που θα υλοποιεί το MLP
    def __init__(self):
        super(MLP, self).__init__() # Ο constructor ορίζει τα στρώματα (layers) και τις παραμέτρους του μοντέλου.
        self.fc1 = nn.Linear(32 * 32 * 3, 1024) # Εισόδος (32x32x3)
        self.fc2 = nn.Linear(1024, 512)          # 1o Κρυφό επίπεδο
        self.fc3 = nn.Linear(512, 256)           # 2o Κρυφό επίπεδο
        self.fc4 = nn.Linear(256,128)            # 3o κρυφό επίπεδο
        self.fc5 = nn.Linear(128,10)             # Έξοδος (10 κλάσεις)
        self.dropout = nn.Dropout(0.4)           # Dropout για γενίκευση

    def forward(self, x):
        x = x.view(-1, 32 * 32 * 3) # Μετατροπή σε μονοδιάστατο διάνυσμα
        x = torch.relu(self.fc1(x))
        x = self.dropout(x)
        x = torch.relu(self.fc2(x))
        x = self.dropout(x)
        x = torch.relu(self.fc3(x))
        x = self.dropout(x)
        x = torch.relu(self.fc4(x))
        x = self.dropout(x)
        x = self.fc5(x)
        return x

# Δημιουργία του μοντέλου
model = MLP().to(device)
```

- **self.fc1 = nn.Linear(32 * 32 * 3, 1024):** Η είσοδος είναι το μέγεθος των δεδομένων CIFAR-10, δηλαδή 32x32 εικόνες με 3 κανάλια χρώματος (RGB). Άρα το συνολικό input είναι $32 * 32 * 3 = 3072$.
- Το στρώμα έχει 1024 νευρώνες, κάτι που το ορίζει ως το πρώτο "κρυφό" επίπεδο (hidden layer) και αντίστοιχα ορίζονται από κάτω το 2° και το 3° .
- **self.fc5 = nn.Linear(128,10):** Έξοδος: 10 νευρώνες, ένας για κάθε κατηγορία της CIFAR-10 (10 κλάσεις).
- **self.dropout = nn.Dropout(0.4):** Το dropout είναι μια τεχνική regularization. Τυχαία "απενεργοποιεί" το 40% των νευρώνων κατά την εκπαίδευση (με πιθανότητα 0.4), βοηθώντας στην αποφυγή overfitting.
- **def forward(self, x):** Η μέθοδος forward περιγράφει πώς τα δεδομένα περνούν μέσα από το δίκτυο.
- **x = x.view(-1, 32 * 32 * 3):** Η είσοδος έχει σχήμα (batch_size, 3, 32, 32) (εικόνες). Η μέθοδος view την "απλώνει" σε μονοδιάστατο διάνυσμα με μέγεθος 3072 (32x32x3), ώστε να μπορεί να περάσει στο fully connected layer.
- **x = torch.relu(self.fc1(x))::** Πέρασμα από το πρώτο πλήρως συνδεδεμένο στρώμα. Η είσοδος περνά από το fc1. Η ενεργοποίηση γίνεται μέσω της συνάρτησης ReLU (**Rectified Linear Unit**), η οποία είναι: $f(x)=\max(0,x)$.
- **x = self.dropout(x):** Εφαρμόζεται dropout στο output του πρώτου στρώματος.
- **x = self.fc5(x):** Δεν εφαρμόζεται ενεργοποίηση εδώ, επειδή η τελική έξοδος περνά από CrossEntropyLoss στο βήμα της απώλειας.
- **model = MLP().to(device):** Το μοντέλο δημιουργείται από την κλάση MLP και μεταφέρεται στη συσκευή (CPU/GPU) που καθορίζεται από την μεταβλητή device.

```
criterion = nn.CrossEntropyLoss()
optimizer = optim.Adam(model.parameters(), lr=0.0004, weight_decay=0.000001)
```

- **criterion = nn.CrossEntropyLoss():** Ως συνάρτηση απώλειας χρησιμοποιείται η CrossEntropyLoss η οποία χρησιμοποιείται κυριώς για προβλήματα classification (βλέπε cifar-10).
- **optimizer = optim.Adam(model.parameters(), lr=0.0004, weight_decay=0.000001):** Χρησιμοποιείται ο Adam ως βελτιστοποιητής.

- **model.parameters():** Τα βάρη του μοντέλου που θα ενημερωθούν.
- **lr=0.0004:** Ρυθμός μάθησης (learning rate), δηλαδή το βήμα που κάνει ο αλγόριθμος στη βελτιστοποίηση.
- **weight_decay=0.000001:** Μειώνει την πολυπλοκότητα του μοντέλου αποφεύγοντας την υπερπροσαρμογή (L2 regularization).

```
# Συνάρτηση αξιολόγησης
def evaluate_model(model, data_loader):
    model.eval() # Θέτουμε το μοντέλο σε κατάσταση αξιολόγησης
    correct = 0
    total = 0
    all_labels = []
    all_predictions = []

    with torch.no_grad():
        for images, labels in data_loader:
            images, labels = images.to(device), labels.to(device)
            outputs = model(images)
            _, predicted = torch.max(outputs, 1)

            total += labels.size(0)
            correct += (predicted == labels).sum().item()

    # Συλλογή προβλέψεων και πραγματικών ετικετών
    all_labels.extend(labels.cpu().numpy())
    all_predictions.extend(predicted.cpu().numpy())

    # Υπολογισμός metrics
    accuracy = 100 * correct / total
    precision = precision_score(all_labels, all_predictions, average='weighted')
    recall = recall_score(all_labels, all_predictions, average='weighted')
    f1 = f1_score(all_labels, all_predictions, average='weighted')

    return accuracy, precision, recall, f1
```

- **def evaluate_model(model, data_loader):** Η συνάρτηση evaluate_model αξιολογεί την απόδοση του μοντέλου σε ένα δοσμένο σύνολο δεδομένων.
- **model.eval():** Θέτει το μοντέλο σε κατάσταση αξιολόγησης και απενεργοποιεί λειτουργίες όπως το Dropout, το οποίο είναι ενεργό μόνο κατά την εκπαίδευση και το Batch Normalization.

- **with torch.no_grad():** Απενεργοποιεί τον υπολογισμό των gradients για λόγους μνήμης και ταχύτητας. Δεν χρειάζονται gradients στην αξιολόγηση, αφού δεν εκπαιδεύουμε το μοντέλο.
 - **for images, labels in data_loader:** Επαναλαμβάνει κάθε batch εικόνων και ετικετών από το data_loader.
 - **images.to(device), labels.to(device):** Μεταφέρει τα δεδομένα στην επιλεγμένη συσκευή (GPU/CPU).
 - **outputs = model(images):** Το μοντέλο επιστρέφει τις προβλέψεις για τις εικόνες.
 - **torch.max(outputs, 1):** Εντοπίζει την κατηγορία με τη μεγαλύτερη πιθανότητα (argmax) για κάθε εικόνα.
 - **labels.size(0):** Ο αριθμός εικόνων στο batch.
 - **predicted == labels:** Ελέγχει αν οι προβλέψεις είναι ίδιες με τις πραγματικές ετικέτες και επιστρέφει έναν πίνακα (Tensor) με True/False.
 - **sum().item():** Υπολογίζει τον αριθμό των σωστών προβλέψεων.
 - **total και correct:** Μετρούν τον συνολικό αριθμό δεδομένων και τις σωστές προβλέψεις, αντίστοιχα.
 - **all_labels.extend(labels.cpu().numpy())**
- all_predictions.extend(predicted.cpu().numpy()):** Οι πραγματικές ετικέτες (labels) και οι προβλέψεις (predicted) αποθηκεύονται για τον υπολογισμό επιπλέον μετρικών (π.χ., precision, recall, F1).
- **cpu().numpy():** Μεταφέρει τα δεδομένα από τη GPU στη CPU και τα μετατρέπει σε numpy arrays.

```
# Εκπαίδευση του μοντέλου
num_epochs = 30
train_accuracies = []
test_accuracies = []
train_losses = []
test_losses = []
best_test_accuracy = 0
patience = 5
epochs_without_improvement = 0
```

- **num_epochs:** Ορίζει τον αριθμό των εποχών εκπαίδευσης.

- **Λίστες (train_accuracies, test_accuracies, train_losses, test_losses):**
Χρησιμοποιούνται για την αποθήκευση των μετρικών ακρίβειας (accuracy) και απώλειας (loss) ανά εποχή.
- **best_test_accuracy:** Κρατάει την καλύτερη ακρίβεια (accuracy) στο σύνολο δοκιμής. Χρησιμοποιείται για να διαπιστώσουμε αν το μοντέλο βελτιώνεται.
- **patience:** Παράμετρος για early stopping. Αν η ακρίβεια στο σύνολο δοκιμής δεν βελτιωθεί για **patience** εποχές, η εκπαίδευση σταματά.
- **epochs_without_improvement:** Μετράει τον αριθμό των συνεχόμενων εποχών χωρίς βελτίωση.

```

for epoch in range(num_epochs):
    model.train()
    running_loss = 0.0

    for images, labels in train_loader:
        images, labels = images.to(device), labels.to(device)

        optimizer.zero_grad()
        outputs = model(images)
        loss = criterion(outputs, labels)
        loss.backward()
        optimizer.step()

        running_loss += loss.item()

    train_losses.append(running_loss / len(train_loader))

    # Αξιολόγηση
    model.eval()
    test_loss = 0.0
    with torch.no_grad(): # Απενεργοποιείται ο υπολογισμός των gradients
        for images, labels in test_loader:
            images, labels = images.to(device), labels.to(device)
            outputs = model(images)
            loss = criterion(outputs, labels)
            test_loss += loss.item()

    test_losses.append(test_loss / len(test_loader))

    # Υπολογισμός metrics για train και test
    train_acc, train_prec, train_recall, train_f1 = evaluate_model(model, train_loader)
    test_acc, test_prec, test_recall, test_f1 = evaluate_model(model, test_loader)

```

- **model.train():** Θέτει το μοντέλο σε κατάσταση εκπαίδευσης. Λειτουργίες όπως το **Dropout** ενεργοποιούνται (χρησιμοποιούνται μόνο κατά την εκπαίδευση).

- **running_loss:** Αρχικοποιείται για να κρατήσει το συνολικό loss για όλα τα batches στην τρέχουσα εποχή.
- **Δεύτερη for:** Οι εικόνες (images) και οι ετικέτες (labels) μεταφέρονται στη συσκευή (GPU/CPU).
- **optimizer.zero_grad():** Καθαρίζει τους παλιούς υπολογισμούς gradients από την προηγούμενη ενημέρωση.
- **outputs = model(images):** Οι έξοδοι του μοντέλου για τις εισόδους του batch.
- **loss = criterion(outputs, labels):** Υπολογίζεται η απώλεια (loss) για τις προβλέψεις του μοντέλου σε σχέση με τις πραγματικές ετικέτες.
- **loss.backward():** Υπολογίζονται τα gradients για τα βάρη του μοντέλου.
- **optimizer.step():** Ενημερώνονται τα βάρη με βάση τα υπολογισμένα gradients.
- **running_loss += loss.item():** Προσθέτει το loss του τρέχοντος batch στο συνολικό loss της εποχής (running_loss).
- **train_losses.append(running_loss / len(train_loader)):** Υπολογίζεται η μέση απώλεια (loss) για όλα τα batches στην εποχή και αποθηκεύεται στη λίστα train_losses
- **model.eval():** Θέτει το μοντέλο σε κατάσταση αξιολόγησης.
- **Για κάθε batch του test set:** Υπολογίζεται το loss και προστίθεται στο test_loss.
- **test_losses.append(test_loss / len(test_loader)):** Υπολογίζεται η μέση απώλεια (loss) για όλα τα batches στο test set και αποθηκεύεται στη λίστα test_losses.
- **train_acc, train_prec, train_recall, train_f1 = evaluate_model(model, train_loader)**
test_acc, test_prec, test_recall, test_f1 = evaluate_model(model, test_loader): Καλείται η συνάρτηση evaluate_model (αναλύθηκε προηγουμένως) για το σύνολο εκπαίδευσης και δοκιμής.

```
# Οπτικοποίηση των αποτελεσμάτων
plt.figure(figsize=(10, 5))
plt.plot(train_accuracies, label='Train Accuracy')
plt.plot(test_accuracies, label='Test Accuracy')
plt.plot(train_losses, label='Train Loss')
plt.plot(test_losses, label='Test Loss')
plt.title("Accuracy per Epoch")
plt.xlabel("Epochs")
plt.ylabel("Accuracy")
plt.legend()
plt.show()
```

- Χρησιμοποιείται η matplotlib για την οπτικοποίηση των αποτελεσμάτων.

2) Εκπαίδευση-Εκτέλεση-Απόδοση

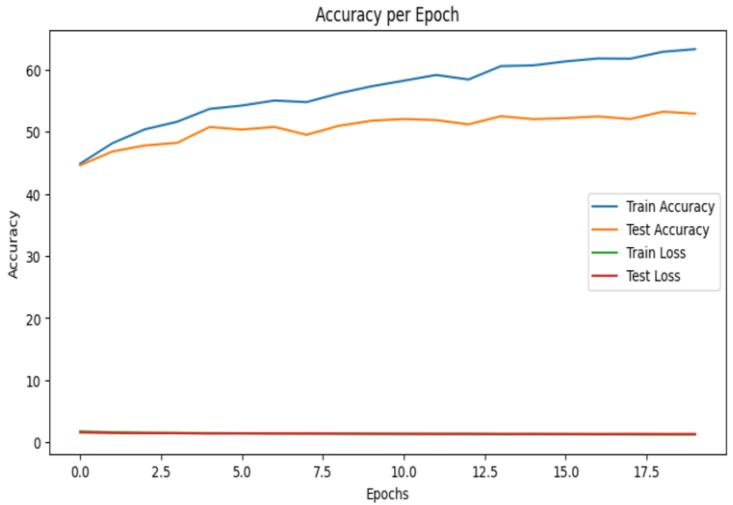
Στο αρχικό state στο νευρωνικό δίκτυο είχαν ανατεθεί οι εξής παράμετροι εκπαίδευσης: Learning Rate=0.001, batch size=50, dropout=0.5, epochs=20, Adam optimizer και ReLu. Από νευρώνες είχε 500 στο πρώτο hidden layer και 250 στο δεύτερο. Χρειάστηκε 22 λεπτά για να εκπαιδευτεί και οι επιδόσεις του ήταν:

θύμας + Κείμενο

```
Train - Loss: 1.3465, Accuracy: 58.23%, Precision: 0.59, Recall: 0.58, F1: 0.58
Test - Loss: 0.2742, Accuracy: 52.07%, Precision: 0.52, Recall: 0.52, F1: 0.52
Epoch [12/20]
Train - Loss: 1.3364, Accuracy: 59.16%, Precision: 0.59, Recall: 0.59, F1: 0.58
Test - Loss: 0.2725, Accuracy: 51.89%, Precision: 0.51, Recall: 0.52, F1: 0.51
Epoch [13/20]
Train - Loss: 1.3210, Accuracy: 58.43%, Precision: 0.60, Recall: 0.58, F1: 0.58
Test - Loss: 0.2755, Accuracy: 51.21%, Precision: 0.53, Recall: 0.51, F1: 0.51
Epoch [14/20]
Train - Loss: 1.3098, Accuracy: 60.59%, Precision: 0.62, Recall: 0.61, F1: 0.60
Test - Loss: 0.2689, Accuracy: 52.52%, Precision: 0.53, Recall: 0.53, F1: 0.52
Epoch [15/20]
Train - Loss: 1.2958, Accuracy: 60.71%, Precision: 0.61, Recall: 0.61, F1: 0.61
Test - Loss: 0.2726, Accuracy: 52.05%, Precision: 0.53, Recall: 0.52, F1: 0.52
Epoch [16/20]
Train - Loss: 1.2883, Accuracy: 61.34%, Precision: 0.62, Recall: 0.61, F1: 0.61
Test - Loss: 0.2702, Accuracy: 52.21%, Precision: 0.53, Recall: 0.52, F1: 0.52
Epoch [17/20]
Train - Loss: 1.2757, Accuracy: 61.82%, Precision: 0.64, Recall: 0.62, F1: 0.62
Test - Loss: 0.2682, Accuracy: 52.48%, Precision: 0.54, Recall: 0.52, F1: 0.52
Epoch [18/20]
Train - Loss: 1.2673, Accuracy: 61.78%, Precision: 0.63, Recall: 0.62, F1: 0.62
Test - Loss: 0.2717, Accuracy: 52.07%, Precision: 0.53, Recall: 0.52, F1: 0.52
Epoch [19/20]
Train - Loss: 1.2501, Accuracy: 62.89%, Precision: 0.64, Recall: 0.63, F1: 0.63
Test - Loss: 0.2680, Accuracy: 53.24%, Precision: 0.54, Recall: 0.53, F1: 0.53
Epoch [20/20]
Train - Loss: 1.2432, Accuracy: 63.32%, Precision: 0.64, Recall: 0.63, F1: 0.63
Test - Loss: 0.2685, Accuracy: 52.93%, Precision: 0.54, Recall: 0.53, F1: 0.53
```

Final Evaluation:

```
Training Set - Accuracy: 63.32%, Precision: 0.64, Recall: 0.63, F1: 0.63
Test Set - Accuracy: 52.93%, Precision: 0.54, Recall: 0.53, F1: 0.53
```



Παρατηρείται ότι το training loss μειώνεται σταδιακά και σταθερά, κάτι που δείχνει ότι το μοντέλο μαθαίνει από το training set. Το training accuracy φτάνει το 63.32%, που είναι αξιοπρεπές αλλά όχι εξαιρετικό για τον αριθμό των 10 κλάσεων. Το test accuracy φτάνει το 52.93% που είναι αρκετά χαμηλότερο από το training accuracy (63.32%). Αυτό μπορεί να σημαίνει ότι το μοντέλο δεν γενικεύει καλά. Το learning rate (0.001) φαίνεται σταθερό, καθώς το training loss μειώνεται σταθερά. Το Dropout (0.5) βοηθά στην αποφυγή overfitting, αλλά αν το μοντέλο είναι ήδη "αδύναμο", μπορεί να περιορίζει τις δυνατότητές του, όπως στη συγκεκριμένη περίπτωση.

Για το επόμενο πείραμα, οι παράμετροι εκπαίδευσης που αλλάζουν είναι μείωση του dropout σε 0.3 και η προσθήκη ενός ακόμα hidden layer και έτσι έχουμε: 800 νευρώνες στο πρώτο hidden layer, 400 στο δεύτερο και 200 στο τρίτο. Η εκπαίδευση ολοκληρώθηκε σε 28 λεπτά και οι επιδόσεις του μοντέλου οι εξής:

```

Epoch [1/20]
Train - Loss: 1.7917, Accuracy: 43.80%, Precision: 0.44, Recall: 0.44, F1: 0.43
Test - Loss: 1.6103, Accuracy: 43.91%, Precision: 0.44, Recall: 0.44, F1: 0.43
Epoch [2/20]
Train - Loss: 1.6493, Accuracy: 47.18%, Precision: 0.48, Recall: 0.47, F1: 0.47
Test - Loss: 1.5530, Accuracy: 46.03%, Precision: 0.47, Recall: 0.46, F1: 0.46
Epoch [3/20]
Train - Loss: 1.5872, Accuracy: 49.01%, Precision: 0.51, Recall: 0.49, F1: 0.48
Test - Loss: 1.5058, Accuracy: 47.08%, Precision: 0.49, Recall: 0.47, F1: 0.46
Epoch [4/20]
Train - Loss: 1.5514, Accuracy: 50.86%, Precision: 0.51, Recall: 0.51, F1: 0.50
Test - Loss: 1.4652, Accuracy: 48.06%, Precision: 0.48, Recall: 0.48, F1: 0.47
Epoch [5/20]
Train - Loss: 1.5166, Accuracy: 52.93%, Precision: 0.52, Recall: 0.53, F1: 0.52
Test - Loss: 1.4473, Accuracy: 49.88%, Precision: 0.49, Recall: 0.50, F1: 0.49
Epoch [6/20]
Train - Loss: 1.4824, Accuracy: 53.24%, Precision: 0.54, Recall: 0.53, F1: 0.53
Test - Loss: 1.4440, Accuracy: 49.08%, Precision: 0.50, Recall: 0.49, F1: 0.48
Epoch [7/20]
Train - Loss: 1.4555, Accuracy: 55.31%, Precision: 0.55, Recall: 0.55, F1: 0.55
Test - Loss: 1.4273, Accuracy: 50.54%, Precision: 0.50, Recall: 0.51, F1: 0.50
Epoch [8/20]
Train - Loss: 1.4307, Accuracy: 55.19%, Precision: 0.56, Recall: 0.55, F1: 0.54
Test - Loss: 1.4239, Accuracy: 49.49%, Precision: 0.50, Recall: 0.49, F1: 0.49
Epoch [9/20]
Train - Loss: 1.4067, Accuracy: 56.73%, Precision: 0.58, Recall: 0.57, F1: 0.57
Test - Loss: 1.4296, Accuracy: 50.38%, Precision: 0.52, Recall: 0.50, F1: 0.50
Epoch [10/20]
Train - Loss: 1.3837, Accuracy: 58.29%, Precision: 0.59, Recall: 0.58, F1: 0.58
Test - Loss: 1.3841, Accuracy: 52.00%, Precision: 0.52, Recall: 0.52, F1: 0.52
Epoch [11/20]
Train - Loss: 1.3655, Accuracy: 59.28%, Precision: 0.60, Recall: 0.59, F1: 0.59
Test - Loss: 1.3800, Accuracy: 51.74%, Precision: 0.53, Recall: 0.52, F1: 0.52

```

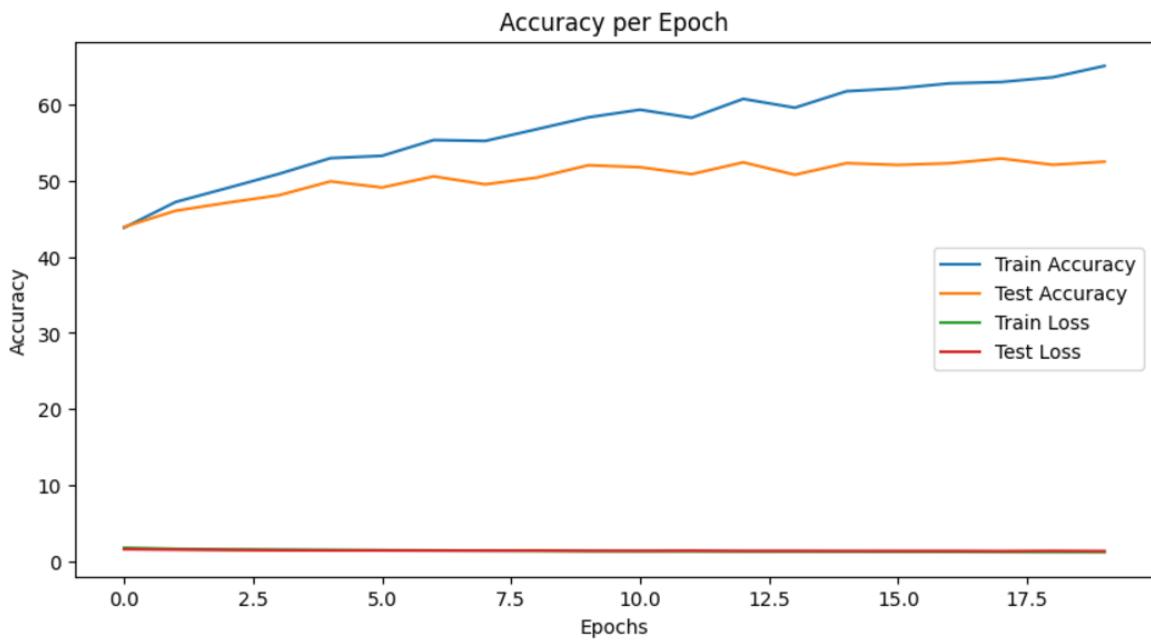
```

Epoch [12/20]
Train - Loss: 1.3457, Accuracy: 58.23%, Precision: 0.59, Recall: 0.58, F1: 0.57
Test - Loss: 1.4103, Accuracy: 50.83%, Precision: 0.52, Recall: 0.51, F1: 0.50
Epoch [13/20]
Train - Loss: 1.3287, Accuracy: 60.70%, Precision: 0.62, Recall: 0.61, F1: 0.61
Test - Loss: 1.3722, Accuracy: 52.38%, Precision: 0.54, Recall: 0.52, F1: 0.52
Epoch [14/20]
Train - Loss: 1.3154, Accuracy: 59.55%, Precision: 0.64, Recall: 0.60, F1: 0.59
Test - Loss: 1.3801, Accuracy: 50.75%, Precision: 0.55, Recall: 0.51, F1: 0.50
Epoch [15/20]
Train - Loss: 1.3058, Accuracy: 61.71%, Precision: 0.63, Recall: 0.62, F1: 0.62
Test - Loss: 1.3641, Accuracy: 52.27%, Precision: 0.53, Recall: 0.52, F1: 0.52
Epoch [16/20]
Train - Loss: 1.2909, Accuracy: 62.09%, Precision: 0.65, Recall: 0.62, F1: 0.62
Test - Loss: 1.3681, Accuracy: 52.04%, Precision: 0.54, Recall: 0.52, F1: 0.52
Epoch [17/20]
Train - Loss: 1.2714, Accuracy: 62.74%, Precision: 0.65, Recall: 0.63, F1: 0.62
Test - Loss: 1.3777, Accuracy: 52.26%, Precision: 0.54, Recall: 0.52, F1: 0.52
Epoch [18/20]
Train - Loss: 1.2521, Accuracy: 62.92%, Precision: 0.64, Recall: 0.63, F1: 0.63
Test - Loss: 1.3602, Accuracy: 52.88%, Precision: 0.54, Recall: 0.53, F1: 0.53
Epoch [19/20]
Train - Loss: 1.2351, Accuracy: 63.54%, Precision: 0.64, Recall: 0.64, F1: 0.64
Test - Loss: 1.3886, Accuracy: 52.06%, Precision: 0.53, Recall: 0.52, F1: 0.52
Epoch [20/20]
Train - Loss: 1.2274, Accuracy: 65.03%, Precision: 0.68, Recall: 0.65, F1: 0.65
Test - Loss: 1.3618, Accuracy: 52.46%, Precision: 0.55, Recall: 0.52, F1: 0.52

```

Final Evaluation:

Training Set - Accuracy: 65.03%, Precision: 0.68, Recall: 0.65, F1: 0.65
Test Set - Accuracy: 52.46%, Precision: 0.55, Recall: 0.52, F1: 0.52

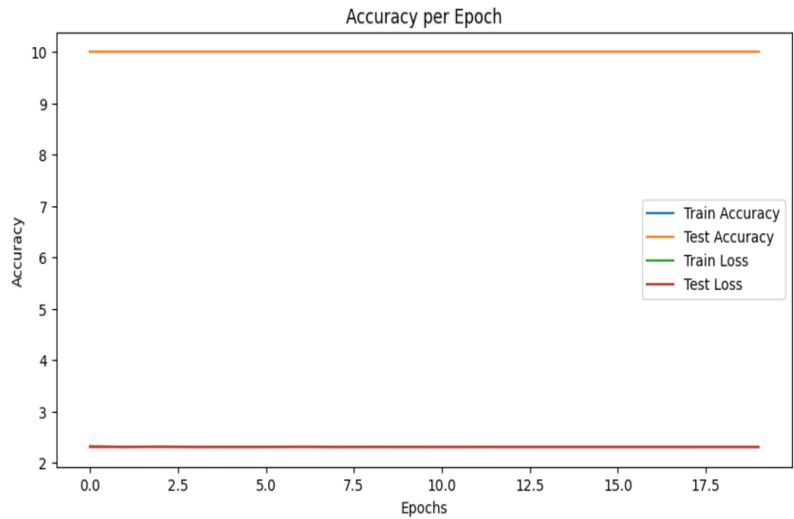


To training accuracy βελτιώνεται συνεχώς και φτάνει στο 65%, ενώ το test accuracy παραμένει κοντά στο 52%. Αυτό δείχνει ότι το μοντέλο μαθαίνει καλύτερα τα δεδομένα εκπαίδευσης, αλλά δεν γενικεύει καλά στα δεδομένα αξιολόγησης.

Στη συνέχεια των πειραμάτων έγινε μία αύξηση του learning rate από 0.001 σε 0.01 ενώ οι υπόλοιπες παράμετροι εκπαίδευσης δεν μεταβλήθηκαν, η εκπαίδευση ολοκληρώθηκε σε 27 λεπτά και τα αποτελέσματα ήταν τα εξής:

```
/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_classification.py:1531: UserWarning: warn_prf(average, modifier, f"{{metric.capitalize()}} is", len(result))
/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_classification.py:1531: UserWarning: warn_prf(average, modifier, f"{{metric.capitalize()}} is", len(result))
Epoch [17/20]
Train - Loss: 2.3038, Accuracy: 10.00%, Precision: 0.01, Recall: 0.10, F1: 0.02
Test - Loss: 2.3035, Accuracy: 10.00%, Precision: 0.01, Recall: 0.10, F1: 0.02
/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_classification.py:1531: UserWarning: warn_prf(average, modifier, f"{{metric.capitalize()}} is", len(result))
/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_classification.py:1531: UserWarning: warn_prf(average, modifier, f"{{metric.capitalize()}} is", len(result))
Epoch [18/20]
Train - Loss: 2.3036, Accuracy: 10.00%, Precision: 0.01, Recall: 0.10, F1: 0.02
Test - Loss: 2.3033, Accuracy: 10.00%, Precision: 0.01, Recall: 0.10, F1: 0.02
/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_classification.py:1531: UserWarning: warn_prf(average, modifier, f"{{metric.capitalize()}} is", len(result))
/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_classification.py:1531: UserWarning: warn_prf(average, modifier, f"{{metric.capitalize()}} is", len(result))
Epoch [19/20]
Train - Loss: 2.3038, Accuracy: 10.00%, Precision: 0.01, Recall: 0.10, F1: 0.02
Test - Loss: 2.3035, Accuracy: 10.00%, Precision: 0.01, Recall: 0.10, F1: 0.02
/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_classification.py:1531: UserWarning: warn_prf(average, modifier, f"{{metric.capitalize()}} is", len(result))
/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_classification.py:1531: UserWarning: warn_prf(average, modifier, f"{{metric.capitalize()}} is", len(result))
Epoch [20/20]
Train - Loss: 2.3035, Accuracy: 10.00%, Precision: 0.01, Recall: 0.10, F1: 0.02
Test - Loss: 2.3039, Accuracy: 10.00%, Precision: 0.01, Recall: 0.10, F1: 0.02
/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_classification.py:1531: UserWarning: warn_prf(average, modifier, f"{{metric.capitalize()}} is", len(result))

Final Evaluation:
Training Set - Accuracy: 10.00%, Precision: 0.01, Recall: 0.10, F1: 0.02
Test Set      - Accuracy: 10.00%, Precision: 0.01, Recall: 0.10, F1: 0.02
```



Το μοντέλο για όλες τις εποχές έχει σταθερά training accuracy=test accuracy=10.00%. Από αυτό συμπεραίνουμε ότι με αυτή την αύξηση του learning rate το μοντέλο δε μαθαίνει καθόλου και λαμβάνει αποφάσεις "στη τύχη". Άρα μία τέτοια απότομη αύξηση του learning rate πρέπει να αποφεύγεται. Όλα δείχνουν πως πρέπει σε επόμενο πείραμα να μειωθεί.

Συνεχίζοντας, το learning rate που έπρεπε να μειωθεί έγινε 0.0005, δηλαδή μισό σε σχέση με το αρχικό και 20 φορές μικρότερο από το προηγούμενο πείραμα. Αυξήσαμε το batch size από 50 σε 70, το dropout από 0.3 σε 0.4 και προσθέσαμε άλλες 10 εποχές. Η εκπαίδευση ολοκληρώθηκε μόλις σε 31 λεπτά με το μοντέλο να αποδίδει:

```

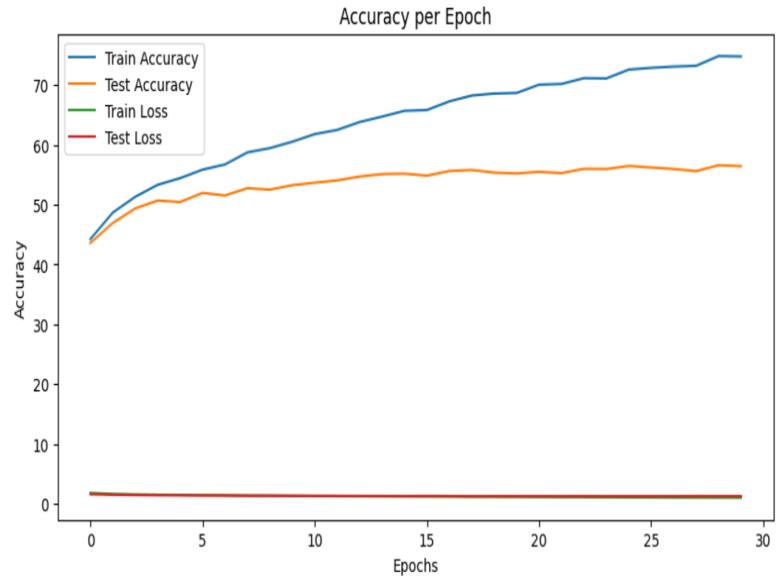
Epoch [1/30]
Train - Loss: 1.8037, Accuracy: 44.23%, Precision: 0.46, Recall: 0.44, F1: 0.43
Test - Loss: 1.5895, Accuracy: 43.58%, Precision: 0.46, Recall: 0.44, F1: 0.42
Epoch [2/30]
Train - Loss: 1.6285, Accuracy: 48.67%, Precision: 0.48, Recall: 0.49, F1: 0.48
Test - Loss: 1.5048, Accuracy: 46.94%, Precision: 0.47, Recall: 0.47, F1: 0.46
Epoch [3/30]
Train - Loss: 1.5539, Accuracy: 51.32%, Precision: 0.52, Recall: 0.51, F1: 0.51
Test - Loss: 1.4562, Accuracy: 49.37%, Precision: 0.50, Recall: 0.49, F1: 0.49
Epoch [4/30]
Train - Loss: 1.5007, Accuracy: 53.32%, Precision: 0.53, Recall: 0.53, F1: 0.53
Test - Loss: 1.4153, Accuracy: 50.69%, Precision: 0.51, Recall: 0.51, F1: 0.50
Epoch [5/30]
Train - Loss: 1.4610, Accuracy: 54.44%, Precision: 0.54, Recall: 0.54, F1: 0.54
Test - Loss: 1.4066, Accuracy: 50.44%, Precision: 0.51, Recall: 0.50, F1: 0.50
Epoch [6/30]
Train - Loss: 1.4264, Accuracy: 55.86%, Precision: 0.56, Recall: 0.56, F1: 0.55
Test - Loss: 1.3791, Accuracy: 51.97%, Precision: 0.52, Recall: 0.52, F1: 0.51
Epoch [7/30]
Train - Loss: 1.3934, Accuracy: 56.73%, Precision: 0.57, Recall: 0.57, F1: 0.56
Test - Loss: 1.3738, Accuracy: 51.54%, Precision: 0.52, Recall: 0.52, F1: 0.51
Epoch [8/30]
Train - Loss: 1.3608, Accuracy: 58.75%, Precision: 0.60, Recall: 0.59, F1: 0.58
Test - Loss: 1.3338, Accuracy: 52.76%, Precision: 0.54, Recall: 0.53, F1: 0.52
Epoch [9/30]
Train - Loss: 1.3312, Accuracy: 59.45%, Precision: 0.60, Recall: 0.59, F1: 0.59
Test - Loss: 1.3435, Accuracy: 52.52%, Precision: 0.53, Recall: 0.53, F1: 0.52
Epoch [10/30]
Train - Loss: 1.3113, Accuracy: 60.53%, Precision: 0.61, Recall: 0.61, F1: 0.60
Test - Loss: 1.3194, Accuracy: 53.26%, Precision: 0.54, Recall: 0.53, F1: 0.53
Epoch [11/30]
Train - Loss: 1.2911, Accuracy: 61.80%, Precision: 0.62, Recall: 0.62, F1: 0.62
Test - Loss: 1.3038, Accuracy: 53.68%, Precision: 0.53, Recall: 0.54, F1: 0.53
Epoch [12/30]

```

```

Epoch [13/30]
Train - Loss: 1.2659, Accuracy: 62.50%, Precision: 0.63, Recall: 0.62, F1: 0.62
Test - Loss: 1.3027, Accuracy: 54.06%, Precision: 0.54, Recall: 0.54, F1: 0.53
Epoch [14/30]
Train - Loss: 1.2480, Accuracy: 63.81%, Precision: 0.64, Recall: 0.64, F1: 0.63
Test - Loss: 1.2918, Accuracy: 54.71%, Precision: 0.55, Recall: 0.55, F1: 0.54
Epoch [15/30]
Train - Loss: 1.2296, Accuracy: 64.72%, Precision: 0.65, Recall: 0.65, F1: 0.64
Test - Loss: 1.2839, Accuracy: 55.12%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [16/30]
Train - Loss: 1.2046, Accuracy: 65.69%, Precision: 0.66, Recall: 0.66, F1: 0.66
Test - Loss: 1.2758, Accuracy: 55.19%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [17/30]
Train - Loss: 1.1932, Accuracy: 65.82%, Precision: 0.66, Recall: 0.66, F1: 0.66
Test - Loss: 1.2921, Accuracy: 54.86%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [18/30]
Train - Loss: 1.1780, Accuracy: 67.28%, Precision: 0.67, Recall: 0.67, F1: 0.67
Test - Loss: 1.2768, Accuracy: 55.63%, Precision: 0.56, Recall: 0.56, F1: 0.55
Epoch [19/30]
Train - Loss: 1.1562, Accuracy: 68.25%, Precision: 0.69, Recall: 0.68, F1: 0.68
Test - Loss: 1.2590, Accuracy: 55.80%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [20/30]
Train - Loss: 1.1422, Accuracy: 68.58%, Precision: 0.69, Recall: 0.69, F1: 0.68
Test - Loss: 1.2710, Accuracy: 55.37%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [21/30]
Train - Loss: 1.1136, Accuracy: 70.07%, Precision: 0.70, Recall: 0.70, F1: 0.70
Test - Loss: 1.2739, Accuracy: 55.50%, Precision: 0.55, Recall: 0.56, F1: 0.55
Epoch [22/30]
Train - Loss: 1.0981, Accuracy: 70.19%, Precision: 0.71, Recall: 0.70, F1: 0.70
Test - Loss: 1.2632, Accuracy: 55.28%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [23/30]

```



Το μοντέλο στις πρώτες εποχές είχε μεγάλες βελτιώσεις τόσο στο training accuracy όσο και στο test accuracy, καθώς και εμφανείς μειώσεις των training loss και test loss. Προς τις τελευταίες εποχές δε συνεχίστηκαν τόσο αισθητές αλλαγές αλλά υπήρχαν κάποιες σταθερές βελτιώσεις, οπότε πολύ πιθανόν αυξάνοντας τις εποχές να έχουμε κάποια καλύτερη απόδοση.

Για το επόμενο πείραμα αυξάνονται οι εποχές από 30 σε 50 και προστίθεται ένα weight decay ίσο με 0.000001 για να αποτρέψουμε υπερεκπαίδευση. Η

εκπαίδευση του μοντέλου ολοκληρώθηκε μετά από 65 λεπτά και τα αποτελέσματα ήταν:

```
Epoch [1/50]
Train - Loss: 1.8023, Accuracy: 45.00%, Precision: 0.45, Recall: 0.45, F1: 0.44
Test - Loss: 1.5891, Accuracy: 44.76%, Precision: 0.44, Recall: 0.45, F1: 0.44
Epoch [2/50]
Train - Loss: 1.6327, Accuracy: 48.17%, Precision: 0.49, Recall: 0.48, F1: 0.47
Test - Loss: 1.5188, Accuracy: 46.74%, Precision: 0.48, Recall: 0.47, F1: 0.46
Epoch [3/50]
Train - Loss: 1.5554, Accuracy: 50.80%, Precision: 0.51, Recall: 0.51, F1: 0.50
Test - Loss: 1.4608, Accuracy: 48.47%, Precision: 0.49, Recall: 0.48, F1: 0.48
Epoch [4/50]
Train - Loss: 1.5041, Accuracy: 52.61%, Precision: 0.53, Recall: 0.53, F1: 0.52
Test - Loss: 1.4173, Accuracy: 50.12%, Precision: 0.51, Recall: 0.50, F1: 0.49
Epoch [5/50]
Train - Loss: 1.4597, Accuracy: 55.51%, Precision: 0.56, Recall: 0.56, F1: 0.55
Test - Loss: 1.3875, Accuracy: 51.56%, Precision: 0.52, Recall: 0.52, F1: 0.51
Epoch [6/50]
Train - Loss: 1.4268, Accuracy: 56.45%, Precision: 0.57, Recall: 0.56, F1: 0.56
Test - Loss: 1.3684, Accuracy: 52.84%, Precision: 0.53, Recall: 0.53, F1: 0.53
Epoch [7/50]
Train - Loss: 1.3894, Accuracy: 58.18%, Precision: 0.59, Recall: 0.58, F1: 0.58
Test - Loss: 1.3675, Accuracy: 53.39%, Precision: 0.54, Recall: 0.53, F1: 0.53
Epoch [8/50]
Train - Loss: 1.3614, Accuracy: 59.09%, Precision: 0.59, Recall: 0.59, F1: 0.59
Test - Loss: 1.3354, Accuracy: 53.18%, Precision: 0.53, Recall: 0.53, F1: 0.53
Epoch [9/50]
Train - Loss: 1.3378, Accuracy: 59.99%, Precision: 0.61, Recall: 0.60, F1: 0.60
Test - Loss: 1.3312, Accuracy: 53.42%, Precision: 0.54, Recall: 0.53, F1: 0.53
Epoch [10/50]
Train - Loss: 1.3097, Accuracy: 60.53%, Precision: 0.61, Recall: 0.61, F1: 0.60
Test - Loss: 1.3141, Accuracy: 53.20%, Precision: 0.53, Recall: 0.53, F1: 0.52
Epoch [11/50]
Train - Loss: 1.2895, Accuracy: 62.26%, Precision: 0.63, Recall: 0.62, F1: 0.62
Test - Loss: 1.3226, Accuracy: 53.67%, Precision: 0.54, Recall: 0.54, F1: 0.53
Epoch [12/50]
```

```
Epoch [23/50]
Train - Loss: 1.0850, Accuracy: 71.08%, Precision: 0.72, Recall: 0.71, F1: 0.71
Test - Loss: 1.2697, Accuracy: 55.89%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [24/50]
Train - Loss: 1.0744, Accuracy: 71.92%, Precision: 0.72, Recall: 0.72, F1: 0.72
Test - Loss: 1.2654, Accuracy: 56.63%, Precision: 0.57, Recall: 0.57, F1: 0.56
Epoch [25/50]
Train - Loss: 1.0616, Accuracy: 71.81%, Precision: 0.73, Recall: 0.72, F1: 0.72
Test - Loss: 1.2684, Accuracy: 56.29%, Precision: 0.57, Recall: 0.56, F1: 0.56
Epoch [26/50]
Train - Loss: 1.0517, Accuracy: 73.01%, Precision: 0.73, Recall: 0.73, F1: 0.73
Test - Loss: 1.2641, Accuracy: 56.23%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [27/50]
Train - Loss: 1.0345, Accuracy: 73.31%, Precision: 0.74, Recall: 0.73, F1: 0.73
Test - Loss: 1.2740, Accuracy: 56.14%, Precision: 0.57, Recall: 0.56, F1: 0.56
Epoch [28/50]
Train - Loss: 1.0220, Accuracy: 74.18%, Precision: 0.74, Recall: 0.74, F1: 0.74
Test - Loss: 1.2713, Accuracy: 56.47%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [29/50]
Train - Loss: 1.0217, Accuracy: 74.57%, Precision: 0.75, Recall: 0.75, F1: 0.75
Test - Loss: 1.2658, Accuracy: 57.11%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [30/50]
Train - Loss: 1.0009, Accuracy: 75.20%, Precision: 0.76, Recall: 0.75, F1: 0.75
Test - Loss: 1.2628, Accuracy: 57.04%, Precision: 0.58, Recall: 0.57, F1: 0.57
Epoch [31/50]
Train - Loss: 0.9869, Accuracy: 75.68%, Precision: 0.76, Recall: 0.76, F1: 0.76
Test - Loss: 1.2535, Accuracy: 57.46%, Precision: 0.58, Recall: 0.57, F1: 0.57
Epoch [32/50]
Train - Loss: 0.9866, Accuracy: 76.14%, Precision: 0.77, Recall: 0.76, F1: 0.76
Test - Loss: 1.2672, Accuracy: 56.43%, Precision: 0.57, Recall: 0.56, F1: 0.57
Epoch [33/50]
Train - Loss: 0.9756, Accuracy: 76.25%, Precision: 0.76, Recall: 0.76, F1: 0.76
Test - Loss: 1.2706, Accuracy: 56.87%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [34/50]
```

```
Epoch [12/50]
Train - Loss: 1.2655, Accuracy: 62.90%, Precision: 0.63, Recall: 0.63, F1: 0.62
Test - Loss: 1.2992, Accuracy: 54.88%, Precision: 0.55, Recall: 0.55, F1: 0.54
Epoch [13/50]
Train - Loss: 1.2452, Accuracy: 64.30%, Precision: 0.65, Recall: 0.64, F1: 0.64
Test - Loss: 1.2927, Accuracy: 54.62%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [14/50]
Train - Loss: 1.2306, Accuracy: 64.58%, Precision: 0.65, Recall: 0.65, F1: 0.65
Test - Loss: 1.2843, Accuracy: 54.91%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [15/50]
Train - Loss: 1.2096, Accuracy: 65.61%, Precision: 0.66, Recall: 0.66, F1: 0.66
Test - Loss: 1.2810, Accuracy: 54.91%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [16/50]
Train - Loss: 1.1881, Accuracy: 66.37%, Precision: 0.67, Recall: 0.66, F1: 0.66
Test - Loss: 1.2825, Accuracy: 55.05%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [17/50]
Train - Loss: 1.1739, Accuracy: 66.84%, Precision: 0.67, Recall: 0.67, F1: 0.67
Test - Loss: 1.2737, Accuracy: 56.23%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [18/50]
Train - Loss: 1.1586, Accuracy: 67.78%, Precision: 0.68, Recall: 0.68, F1: 0.68
Test - Loss: 1.2766, Accuracy: 55.14%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [19/50]
Train - Loss: 1.1455, Accuracy: 69.12%, Precision: 0.69, Recall: 0.69, F1: 0.69
Test - Loss: 1.2632, Accuracy: 55.80%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [20/50]
Train - Loss: 1.1320, Accuracy: 68.63%, Precision: 0.69, Recall: 0.69, F1: 0.68
Test - Loss: 1.2709, Accuracy: 55.76%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [21/50]
Train - Loss: 1.1101, Accuracy: 69.73%, Precision: 0.71, Recall: 0.70, F1: 0.70
Test - Loss: 1.2689, Accuracy: 55.64%, Precision: 0.57, Recall: 0.56, F1: 0.55
Epoch [22/50]
Train - Loss: 1.1050, Accuracy: 70.27%, Precision: 0.70, Recall: 0.70, F1: 0.70
Test - Loss: 1.2671, Accuracy: 55.67%, Precision: 0.56, Recall: 0.56, F1: 0.55
```

```
Train - Loss: 0.9656, Accuracy: 77.08%, Precision: 0.78, Recall: 0.77, F1: 0.77
Test - Loss: 1.2704, Accuracy: 56.74%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [35/50]
Train - Loss: 0.9544, Accuracy: 76.98%, Precision: 0.78, Recall: 0.77, F1: 0.77
Test - Loss: 1.2744, Accuracy: 56.61%, Precision: 0.57, Recall: 0.57, F1: 0.56
Epoch [36/50]
Train - Loss: 0.9484, Accuracy: 77.70%, Precision: 0.78, Recall: 0.78, F1: 0.78
Test - Loss: 1.2600, Accuracy: 57.29%, Precision: 0.58, Recall: 0.57, F1: 0.57
Epoch [37/50]
Train - Loss: 0.9384, Accuracy: 77.96%, Precision: 0.78, Recall: 0.78, F1: 0.78
Test - Loss: 1.2732, Accuracy: 56.39%, Precision: 0.57, Recall: 0.56, F1: 0.56
Epoch [38/50]
Train - Loss: 0.9244, Accuracy: 78.63%, Precision: 0.79, Recall: 0.79, F1: 0.79
Test - Loss: 1.2805, Accuracy: 56.75%, Precision: 0.58, Recall: 0.57, F1: 0.57
Epoch [39/50]
Train - Loss: 0.9242, Accuracy: 78.88%, Precision: 0.80, Recall: 0.79, F1: 0.79
Test - Loss: 1.2690, Accuracy: 56.85%, Precision: 0.58, Recall: 0.57, F1: 0.57
Epoch [40/50]
Train - Loss: 0.9190, Accuracy: 79.36%, Precision: 0.80, Recall: 0.79, F1: 0.79
Test - Loss: 1.2782, Accuracy: 56.83%, Precision: 0.58, Recall: 0.57, F1: 0.57
Epoch [41/50]
Train - Loss: 0.8970, Accuracy: 79.86%, Precision: 0.80, Recall: 0.80, F1: 0.80
Test - Loss: 1.2842, Accuracy: 57.32%, Precision: 0.58, Recall: 0.57, F1: 0.57
Epoch [42/50]
Train - Loss: 0.8964, Accuracy: 79.47%, Precision: 0.81, Recall: 0.79, F1: 0.80
Test - Loss: 1.2852, Accuracy: 56.48%, Precision: 0.58, Recall: 0.56, F1: 0.57
Epoch [43/50]
Train - Loss: 0.8911, Accuracy: 80.23%, Precision: 0.81, Recall: 0.80, F1: 0.80
Test - Loss: 1.2847, Accuracy: 56.96%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [44/50]
Train - Loss: 0.8751, Accuracy: 79.87%, Precision: 0.80, Recall: 0.80, F1: 0.80
Test - Loss: 1.2959, Accuracy: 56.24%, Precision: 0.57, Recall: 0.56, F1: 0.56
```

```

Train - Loss: 0.8970, Accuracy: 79.86%, Precision: 0.80, Recall: 0.80, F1: 0.80
Test - Loss: 1.2842, Accuracy: 57.32%, Precision: 0.58, Recall: 0.57, F1: 0.57
Epoch [42/50]
Train - Loss: 0.8964, Accuracy: 79.47%, Precision: 0.81, Recall: 0.79, F1: 0.80
Test - Loss: 1.2852, Accuracy: 56.48%, Precision: 0.58, Recall: 0.56, F1: 0.57
Epoch [43/50]
Train - Loss: 0.8911, Accuracy: 80.23%, Precision: 0.81, Recall: 0.80, F1: 0.80
Test - Loss: 1.2847, Accuracy: 56.96%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [44/50]
Train - Loss: 0.8751, Accuracy: 79.87%, Precision: 0.80, Recall: 0.80, F1: 0.80
Test - Loss: 1.2959, Accuracy: 56.24%, Precision: 0.57, Recall: 0.56, F1: 0.56
Epoch [45/50]
Train - Loss: 0.8735, Accuracy: 81.12%, Precision: 0.81, Recall: 0.81, F1: 0.81
Test - Loss: 1.2885, Accuracy: 56.81%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [46/50]
Train - Loss: 0.8647, Accuracy: 81.26%, Precision: 0.82, Recall: 0.81, F1: 0.81
Test - Loss: 1.2911, Accuracy: 56.96%, Precision: 0.58, Recall: 0.57, F1: 0.57
Epoch [47/50]
Train - Loss: 0.8579, Accuracy: 81.66%, Precision: 0.82, Recall: 0.82, F1: 0.82
Test - Loss: 1.3018, Accuracy: 56.97%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [48/50]
Train - Loss: 0.8482, Accuracy: 82.02%, Precision: 0.82, Recall: 0.82, F1: 0.82
Test - Loss: 1.3004, Accuracy: 57.28%, Precision: 0.58, Recall: 0.57, F1: 0.57
Epoch [49/50]
Train - Loss: 0.8544, Accuracy: 81.97%, Precision: 0.82, Recall: 0.82, F1: 0.82
Test - Loss: 1.3068, Accuracy: 56.87%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [50/50]
Train - Loss: 0.8437, Accuracy: 82.62%, Precision: 0.83, Recall: 0.83, F1: 0.83
Test - Loss: 1.2879, Accuracy: 57.70%, Precision: 0.58, Recall: 0.58, F1: 0.58

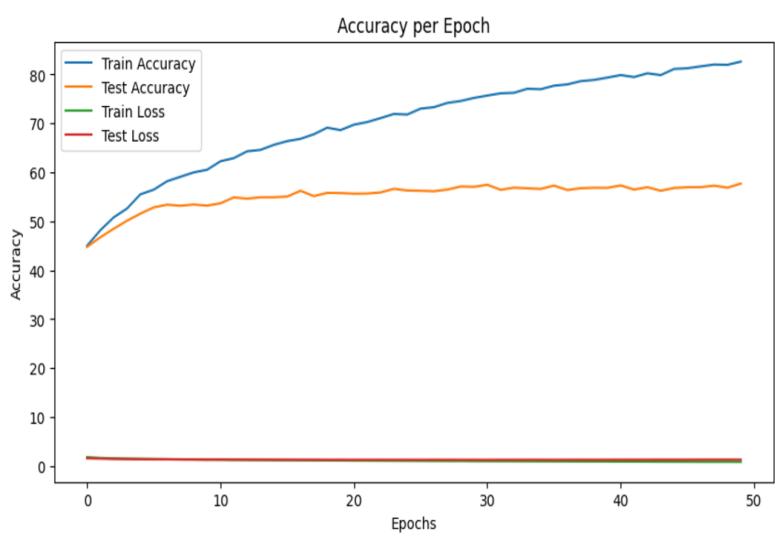
```

Final Evaluation:

```

Training Set - Accuracy: 82.62%, Precision: 0.83, Recall: 0.83, F1: 0.83
Test Set      - Accuracy: 57.70%, Precision: 0.58, Recall: 0.58, F1: 0.58

```



Όπως φαίνεται, πετύχαμε καλή βελτίωση στην εκπαίδευση (train accuracy=82.62%) αλλά το test accuracy είχε μικρότερη βελτίωση (περίπου 1%). Επίσης, μετά την 31^η εποχή, το test accuracy είχε σταματήσει να βελτιώνεται (και σε κάποιες εποχές μάλιστα μειωνόταν ελάχιστα). Αυτό σε συνδυασμό με τη μεγάλη διαφορά που είχε από το train accuracy, αποτελούν σημάδια υπερεκπαίδευσης. Σε αυτή τη περίπτωση, ίσως πρέπει να μειωθούν εποχές ή να αυξηθεί το regularization (μεγαλύτερο dropout ή πιο δυναμικό weight decay).

Με βάση τα παραπάνω συμπεράσματα, για το επόμενο πείραμα θα αυξήσουμε το dropout από 0.4 σε 0.5, θα μειώσουμε τις εποχές από 50 σε 40 και θα μειώσουμε το weight decay σε 0.00001. Επίσης, θα αυξήσουμε το batch size από 70 σε 120 αποσκοπώντας σε πιο σταθερές εκτιμήσεις του gradient. Οι υπόλοιπες παράμετροι εκπαίδευσης παραμένουν ίδιες (learning rate=0.0005, Adam optimizer, ReLu και νευρώνες 800 στο πρώτο hidden layer, 400 στο δεύτερο και 200 στο τρίτο). Η εκπαίδευση ολοκληρώθηκε σε 39 λεπτά και η απόδοση ήταν:

```

Epoch [1/40]
Train - Loss: 1.8735, Accuracy: 42.98%, Precision: 0.43, Recall: 0.43, F1: 0.42
Test - Loss: 1.6306, Accuracy: 42.74%, Precision: 0.42, Recall: 0.43, F1: 0.42
Epoch [2/40]
Train - Loss: 1.6993, Accuracy: 46.12%, Precision: 0.46, Recall: 0.46, F1: 0.45
Test - Loss: 1.5564, Accuracy: 45.40%, Precision: 0.46, Recall: 0.45, F1: 0.44
Epoch [3/40]
Train - Loss: 1.6262, Accuracy: 49.19%, Precision: 0.49, Recall: 0.49, F1: 0.49
Test - Loss: 1.5108, Accuracy: 48.09%, Precision: 0.48, Recall: 0.48, F1: 0.48
Epoch [4/40]
Train - Loss: 1.5778, Accuracy: 50.35%, Precision: 0.51, Recall: 0.50, F1: 0.49
Test - Loss: 1.4654, Accuracy: 48.40%, Precision: 0.48, Recall: 0.48, F1: 0.47
Epoch [5/40]
Train - Loss: 1.5380, Accuracy: 52.04%, Precision: 0.52, Recall: 0.52, F1: 0.51
Test - Loss: 1.4452, Accuracy: 49.94%, Precision: 0.49, Recall: 0.50, F1: 0.49
Epoch [6/40]
Train - Loss: 1.5037, Accuracy: 53.48%, Precision: 0.53, Recall: 0.53, F1: 0.53
Test - Loss: 1.4189, Accuracy: 50.64%, Precision: 0.50, Recall: 0.51, F1: 0.50
Epoch [7/40]
Train - Loss: 1.4778, Accuracy: 54.47%, Precision: 0.54, Recall: 0.54, F1: 0.54
Test - Loss: 1.4075, Accuracy: 50.99%, Precision: 0.51, Recall: 0.51, F1: 0.50
Epoch [8/40]
Train - Loss: 1.4577, Accuracy: 55.62%, Precision: 0.55, Recall: 0.56, F1: 0.55
Test - Loss: 1.3848, Accuracy: 51.82%, Precision: 0.51, Recall: 0.52, F1: 0.51
Epoch [9/40]
Train - Loss: 1.4329, Accuracy: 56.32%, Precision: 0.56, Recall: 0.56, F1: 0.56
Test - Loss: 1.3745, Accuracy: 52.18%, Precision: 0.52, Recall: 0.52, F1: 0.52
Epoch [10/40]
Train - Loss: 1.4134, Accuracy: 56.99%, Precision: 0.57, Recall: 0.57, F1: 0.57
Test - Loss: 1.3527, Accuracy: 52.64%, Precision: 0.53, Recall: 0.53, F1: 0.52
Epoch [11/40]
Train - Loss: 1.3955, Accuracy: 58.53%, Precision: 0.58, Recall: 0.59, F1: 0.58
Test - Loss: 1.3433, Accuracy: 53.66%, Precision: 0.53, Recall: 0.54, F1: 0.53
Epoch [12/40]
Train - Loss: 1.3736, Accuracy: 58.72%, Precision: 0.59, Recall: 0.59, F1: 0.58
Test - Loss: 1.3401, Accuracy: 53.19%, Precision: 0.53, Recall: 0.53, F1: 0.52
Epoch [13/40]
Train - Loss: 1.3619, Accuracy: 59.74%, Precision: 0.60, Recall: 0.60, F1: 0.59
Test - Loss: 1.3303, Accuracy: 53.52%, Precision: 0.54, Recall: 0.54, F1: 0.53
Epoch [14/40]
Train - Loss: 1.3386, Accuracy: 60.17%, Precision: 0.60, Recall: 0.60, F1: 0.60
Test - Loss: 1.3337, Accuracy: 53.59%, Precision: 0.53, Recall: 0.54, F1: 0.53
Epoch [15/40]
Train - Loss: 1.3285, Accuracy: 60.62%, Precision: 0.61, Recall: 0.61, F1: 0.60
Test - Loss: 1.3173, Accuracy: 54.12%, Precision: 0.54, Recall: 0.54, F1: 0.54
Epoch [16/40]
Train - Loss: 1.3146, Accuracy: 61.77%, Precision: 0.62, Recall: 0.62, F1: 0.62
Test - Loss: 1.3111, Accuracy: 54.41%, Precision: 0.54, Recall: 0.54, F1: 0.54
Epoch [17/40]
Train - Loss: 1.2969, Accuracy: 62.96%, Precision: 0.63, Recall: 0.63, F1: 0.63
Test - Loss: 1.3085, Accuracy: 54.79%, Precision: 0.54, Recall: 0.55, F1: 0.55
Epoch [18/40]
Train - Loss: 1.2866, Accuracy: 62.82%, Precision: 0.64, Recall: 0.63, F1: 0.62
Test - Loss: 1.3059, Accuracy: 54.46%, Precision: 0.55, Recall: 0.54, F1: 0.54
Epoch [19/40]
Train - Loss: 1.2735, Accuracy: 63.56%, Precision: 0.64, Recall: 0.64, F1: 0.63
Test - Loss: 1.2911, Accuracy: 54.61%, Precision: 0.55, Recall: 0.55, F1: 0.54
Epoch [20/40]
Train - Loss: 1.2615, Accuracy: 64.55%, Precision: 0.65, Recall: 0.65, F1: 0.64
Test - Loss: 1.2839, Accuracy: 55.42%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [21/40]
Train - Loss: 1.2468, Accuracy: 64.51%, Precision: 0.64, Recall: 0.65, F1: 0.64
Test - Loss: 1.2940, Accuracy: 55.22%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [22/40]
Train - Loss: 1.2364, Accuracy: 65.44%, Precision: 0.65, Recall: 0.65, F1: 0.65
Test - Loss: 1.2845, Accuracy: 55.67%, Precision: 0.55, Recall: 0.56, F1: 0.55
Epoch [23/40]
Train - Loss: 1.2260, Accuracy: 66.29%, Precision: 0.67, Recall: 0.66, F1: 0.66

```

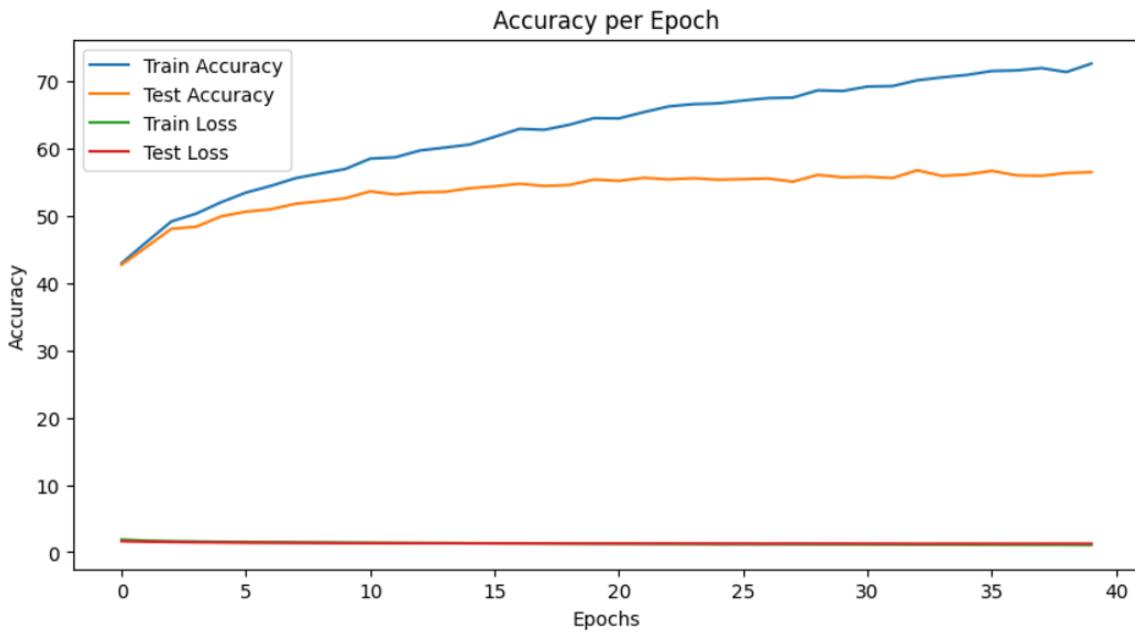
```

Epoch [23/40]
Train - Loss: 1.2260, Accuracy: 66.29%, Precision: 0.67, Recall: 0.66, F1: 0.66
Test - Loss: 1.2896, Accuracy: 55.44%, Precision: 0.56, Recall: 0.55, F1: 0.56
Epoch [24/40]
Train - Loss: 1.2146, Accuracy: 66.63%, Precision: 0.67, Recall: 0.67, F1: 0.67
Test - Loss: 1.2803, Accuracy: 55.61%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [25/40]
Train - Loss: 1.1996, Accuracy: 66.75%, Precision: 0.68, Recall: 0.67, F1: 0.67
Test - Loss: 1.2710, Accuracy: 55.39%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [26/40]
Train - Loss: 1.1949, Accuracy: 67.17%, Precision: 0.68, Recall: 0.67, F1: 0.67
Test - Loss: 1.2744, Accuracy: 55.47%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [27/40]
Train - Loss: 1.1833, Accuracy: 67.52%, Precision: 0.67, Recall: 0.68, F1: 0.67
Test - Loss: 1.2621, Accuracy: 55.57%, Precision: 0.55, Recall: 0.56, F1: 0.55
Epoch [28/40]
Train - Loss: 1.1791, Accuracy: 67.60%, Precision: 0.69, Recall: 0.68, F1: 0.67
Test - Loss: 1.2769, Accuracy: 55.11%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [29/40]
Train - Loss: 1.1702, Accuracy: 68.68%, Precision: 0.69, Recall: 0.69, F1: 0.69
Test - Loss: 1.2705, Accuracy: 56.11%, Precision: 0.57, Recall: 0.56, F1: 0.56
Epoch [30/40]
Train - Loss: 1.1620, Accuracy: 68.58%, Precision: 0.69, Recall: 0.69, F1: 0.69
Test - Loss: 1.2740, Accuracy: 55.75%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [31/40]
Train - Loss: 1.1529, Accuracy: 69.25%, Precision: 0.69, Recall: 0.69, F1: 0.69
Test - Loss: 1.2622, Accuracy: 55.84%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [32/40]
Train - Loss: 1.1368, Accuracy: 69.31%, Precision: 0.71, Recall: 0.69, F1: 0.69
Test - Loss: 1.2692, Accuracy: 55.64%, Precision: 0.57, Recall: 0.56, F1: 0.55
Epoch [33/40]
Train - Loss: 1.1302, Accuracy: 70.18%, Precision: 0.70, Recall: 0.70, F1: 0.70
Test - Loss: 1.2547, Accuracy: 56.79%, Precision: 0.57, Recall: 0.57, F1: 0.56
Epoch [34/40]
Train - Loss: 1.1232, Accuracy: 70.61%, Precision: 0.71, Recall: 0.71, F1: 0.71
Test - Loss: 1.2618, Accuracy: 55.96%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [35/40]
Train - Loss: 1.1142, Accuracy: 70.98%, Precision: 0.72, Recall: 0.71, F1: 0.71
Test - Loss: 1.2686, Accuracy: 56.17%, Precision: 0.57, Recall: 0.56, F1: 0.56
Epoch [36/40]
Train - Loss: 1.1069, Accuracy: 71.56%, Precision: 0.72, Recall: 0.72, F1: 0.71
Test - Loss: 1.2611, Accuracy: 56.71%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [37/40]
Train - Loss: 1.0991, Accuracy: 71.65%, Precision: 0.72, Recall: 0.72, F1: 0.72
Test - Loss: 1.2658, Accuracy: 56.04%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [38/40]
Train - Loss: 1.0938, Accuracy: 71.99%, Precision: 0.73, Recall: 0.72, F1: 0.72
Test - Loss: 1.2656, Accuracy: 55.97%, Precision: 0.57, Recall: 0.56, F1: 0.56
Epoch [39/40]
Train - Loss: 1.0865, Accuracy: 71.40%, Precision: 0.72, Recall: 0.71, F1: 0.71
Test - Loss: 1.2685, Accuracy: 56.39%, Precision: 0.57, Recall: 0.56, F1: 0.56
Epoch [40/40]
Train - Loss: 1.0727, Accuracy: 72.66%, Precision: 0.74, Recall: 0.73, F1: 0.73
Test - Loss: 1.2626, Accuracy: 56.51%, Precision: 0.57, Recall: 0.57, F1: 0.57

```

Final Evaluation:

Training Set - Accuracy: 72.66%, Precision: 0.74, Recall: 0.73, F1: 0.73
 Test Set - Accuracy: 56.51%, Precision: 0.57, Recall: 0.57, F1: 0.57



Παρατηρείται μείωση 10% στο training accuracy, με τα υπόλοιπα accuracies και losses να μην αλλάζουν αισθητά. Αυτή η αλλαγή ίσως οφείλεται στη μεγάλη αύξηση του batch size και λόγω αυτού ίσως να χανόταν πληροφορία κατά την εκπαίδευση. Στη συνέχεια θα το μειώσουμε σε 64. Η εκπαίδευση ολοκληρώθηκε σε 50 λεπτά.

```

Epoch [1/40]
Train - Loss: 1.8655, Accuracy: 42.53%, Precision: 0.43, Recall: 0.43, F1: 0.42
Test - Loss: 1.6419, Accuracy: 42.41%, Precision: 0.43, Recall: 0.42, F1: 0.41
Epoch [2/40]
Train - Loss: 1.7060, Accuracy: 45.96%, Precision: 0.46, Recall: 0.46, F1: 0.45
Test - Loss: 1.5636, Accuracy: 45.80%, Precision: 0.46, Recall: 0.46, F1: 0.45
Epoch [3/40]
Train - Loss: 1.6376, Accuracy: 48.53%, Precision: 0.48, Recall: 0.49, F1: 0.47
Test - Loss: 1.5082, Accuracy: 47.82%, Precision: 0.47, Recall: 0.48, F1: 0.46
Epoch [4/40]
Train - Loss: 1.5994, Accuracy: 50.54%, Precision: 0.51, Recall: 0.51, F1: 0.50
Test - Loss: 1.4841, Accuracy: 49.07%, Precision: 0.49, Recall: 0.49, F1: 0.48
Epoch [5/40]
Train - Loss: 1.5635, Accuracy: 51.17%, Precision: 0.51, Recall: 0.51, F1: 0.51
Test - Loss: 1.4752, Accuracy: 48.49%, Precision: 0.49, Recall: 0.48, F1: 0.48
Epoch [6/40]
Train - Loss: 1.5425, Accuracy: 52.26%, Precision: 0.52, Recall: 0.52, F1: 0.52
Test - Loss: 1.4483, Accuracy: 49.87%, Precision: 0.50, Recall: 0.50, F1: 0.49
Epoch [7/40]
Train - Loss: 1.5119, Accuracy: 53.88%, Precision: 0.54, Recall: 0.54, F1: 0.53
Test - Loss: 1.4258, Accuracy: 50.32%, Precision: 0.50, Recall: 0.50, F1: 0.50
Epoch [8/40]
Train - Loss: 1.4862, Accuracy: 54.56%, Precision: 0.54, Recall: 0.55, F1: 0.54
Test - Loss: 1.4108, Accuracy: 51.31%, Precision: 0.51, Recall: 0.51, F1: 0.51
Epoch [9/40]
Train - Loss: 1.4718, Accuracy: 55.05%, Precision: 0.56, Recall: 0.55, F1: 0.55
Test - Loss: 1.4084, Accuracy: 50.78%, Precision: 0.52, Recall: 0.51, F1: 0.50
Epoch [10/40]
Train - Loss: 1.4473, Accuracy: 56.47%, Precision: 0.57, Recall: 0.56, F1: 0.56
Test - Loss: 1.3816, Accuracy: 51.99%, Precision: 0.52, Recall: 0.52, F1: 0.52
Epoch [11/40]
Train - Loss: 1.4324, Accuracy: 57.00%, Precision: 0.58, Recall: 0.57, F1: 0.57
Test - Loss: 1.3641, Accuracy: 52.13%, Precision: 0.53, Recall: 0.52, F1: 0.52
Epoch [12/40]
Train - Loss: 1.4131, Accuracy: 57.45%, Precision: 0.59, Recall: 0.57, F1: 0.57

```

```

Epoch [12/40]
Train - Loss: 1.4131, Accuracy: 57.45%, Precision: 0.58, Recall: 0.57, F1: 0.57
Test - Loss: 1.3775, Accuracy: 52.23%, Precision: 0.52, Recall: 0.52, F1: 0.52
Epoch [13/40]
Train - Loss: 1.4020, Accuracy: 58.25%, Precision: 0.59, Recall: 0.58, F1: 0.58
Test - Loss: 1.3703, Accuracy: 53.06%, Precision: 0.53, Recall: 0.53, F1: 0.53
Epoch [14/40]
Train - Loss: 1.3834, Accuracy: 58.65%, Precision: 0.59, Recall: 0.59, F1: 0.58
Test - Loss: 1.3536, Accuracy: 52.66%, Precision: 0.53, Recall: 0.53, F1: 0.52
Epoch [15/40]
Train - Loss: 1.3793, Accuracy: 59.56%, Precision: 0.60, Recall: 0.60, F1: 0.59
Test - Loss: 1.3483, Accuracy: 53.00%, Precision: 0.54, Recall: 0.53, F1: 0.53
Epoch [16/40]
Train - Loss: 1.3583, Accuracy: 59.27%, Precision: 0.60, Recall: 0.59, F1: 0.59
Test - Loss: 1.3477, Accuracy: 52.98%, Precision: 0.53, Recall: 0.53, F1: 0.53
Epoch [17/40]
Train - Loss: 1.3553, Accuracy: 60.52%, Precision: 0.61, Recall: 0.61, F1: 0.60
Test - Loss: 1.3290, Accuracy: 53.82%, Precision: 0.55, Recall: 0.54, F1: 0.53
Epoch [18/40]
Train - Loss: 1.3355, Accuracy: 61.16%, Precision: 0.61, Recall: 0.61, F1: 0.61
Test - Loss: 1.3393, Accuracy: 53.26%, Precision: 0.53, Recall: 0.53, F1: 0.53
Epoch [19/40]
Train - Loss: 1.3275, Accuracy: 61.12%, Precision: 0.62, Recall: 0.61, F1: 0.61
Test - Loss: 1.3352, Accuracy: 52.89%, Precision: 0.54, Recall: 0.53, F1: 0.52
Epoch [20/40]
Train - Loss: 1.3144, Accuracy: 61.68%, Precision: 0.63, Recall: 0.62, F1: 0.62
Test - Loss: 1.3425, Accuracy: 53.58%, Precision: 0.55, Recall: 0.54, F1: 0.53
Epoch [21/40]
Train - Loss: 1.3031, Accuracy: 61.45%, Precision: 0.63, Recall: 0.61, F1: 0.61
Test - Loss: 1.3441, Accuracy: 53.25%, Precision: 0.54, Recall: 0.53, F1: 0.53
Epoch [22/40]
Train - Loss: 1.2979, Accuracy: 63.51%, Precision: 0.64, Recall: 0.64, F1: 0.63
Test - Loss: 1.3164, Accuracy: 54.52%, Precision: 0.55, Recall: 0.55, F1: 0.54
Epoch [23/40]
Train - Loss: 1.2805, Accuracy: 63.96%, Precision: 0.65, Recall: 0.64, F1: 0.64

```

```

Train - Loss: 1.2805, Accuracy: 63.96%, Precision: 0.65, Recall: 0.64, F1: 0.64
Test - Loss: 1.3104, Accuracy: 54.70%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [24/40]
Train - Loss: 1.2794, Accuracy: 64.13%, Precision: 0.65, Recall: 0.64, F1: 0.64
Test - Loss: 1.3128, Accuracy: 54.22%, Precision: 0.55, Recall: 0.54, F1: 0.54
Epoch [25/40]
Train - Loss: 1.2681, Accuracy: 64.35%, Precision: 0.65, Recall: 0.64, F1: 0.64
Test - Loss: 1.3155, Accuracy: 54.53%, Precision: 0.55, Recall: 0.55, F1: 0.54
Epoch [26/40]
Train - Loss: 1.2615, Accuracy: 65.11%, Precision: 0.66, Recall: 0.65, F1: 0.65
Test - Loss: 1.3050, Accuracy: 54.99%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [27/40]
Train - Loss: 1.2461, Accuracy: 64.47%, Precision: 0.66, Recall: 0.64, F1: 0.64
Test - Loss: 1.3081, Accuracy: 54.28%, Precision: 0.55, Recall: 0.54, F1: 0.54
Epoch [28/40]
Train - Loss: 1.2477, Accuracy: 64.95%, Precision: 0.66, Recall: 0.65, F1: 0.65
Test - Loss: 1.3070, Accuracy: 54.23%, Precision: 0.55, Recall: 0.54, F1: 0.54
Epoch [29/40]
Train - Loss: 1.2322, Accuracy: 66.09%, Precision: 0.67, Recall: 0.66, F1: 0.66
Test - Loss: 1.3072, Accuracy: 55.00%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [30/40]
Train - Loss: 1.2256, Accuracy: 66.06%, Precision: 0.68, Recall: 0.66, F1: 0.66
Test - Loss: 1.3099, Accuracy: 54.67%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [31/40]
Train - Loss: 1.2146, Accuracy: 67.18%, Precision: 0.68, Recall: 0.67, F1: 0.67
Test - Loss: 1.2987, Accuracy: 55.11%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [32/40]
Train - Loss: 1.2109, Accuracy: 67.41%, Precision: 0.68, Recall: 0.67, F1: 0.67
Test - Loss: 1.2981, Accuracy: 55.12%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [33/40]
Train - Loss: 1.2052, Accuracy: 67.80%, Precision: 0.69, Recall: 0.68, F1: 0.68
Test - Loss: 1.2866, Accuracy: 55.29%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [34/40]
Train - Loss: 1.1949, Accuracy: 67.60%, Precision: 0.68, Recall: 0.68, F1: 0.67
Test - Loss: 1.2943, Accuracy: 55.27%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [35/40]
Train - Loss: 1.1914, Accuracy: 68.00%, Precision: 0.69, Recall: 0.68, F1: 0.68
Test - Loss: 1.2915, Accuracy: 55.26%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [36/40]
Train - Loss: 1.1879, Accuracy: 68.90%, Precision: 0.70, Recall: 0.69, F1: 0.69
Test - Loss: 1.2874, Accuracy: 55.39%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [37/40]
Train - Loss: 1.1782, Accuracy: 68.40%, Precision: 0.69, Recall: 0.68, F1: 0.69
Test - Loss: 1.2863, Accuracy: 55.10%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [38/40]
Train - Loss: 1.1777, Accuracy: 69.20%, Precision: 0.70, Recall: 0.69, F1: 0.69
Test - Loss: 1.2847, Accuracy: 55.77%, Precision: 0.57, Recall: 0.56, F1: 0.56
Epoch [39/40]
Train - Loss: 1.1654, Accuracy: 69.24%, Precision: 0.71, Recall: 0.69, F1: 0.69
Test - Loss: 1.2977, Accuracy: 55.03%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [40/40]
Train - Loss: 1.1584, Accuracy: 70.03%, Precision: 0.70, Recall: 0.70, F1: 0.70
Test - Loss: 1.2958, Accuracy: 55.69%, Precision: 0.56, Recall: 0.56, F1: 0.56

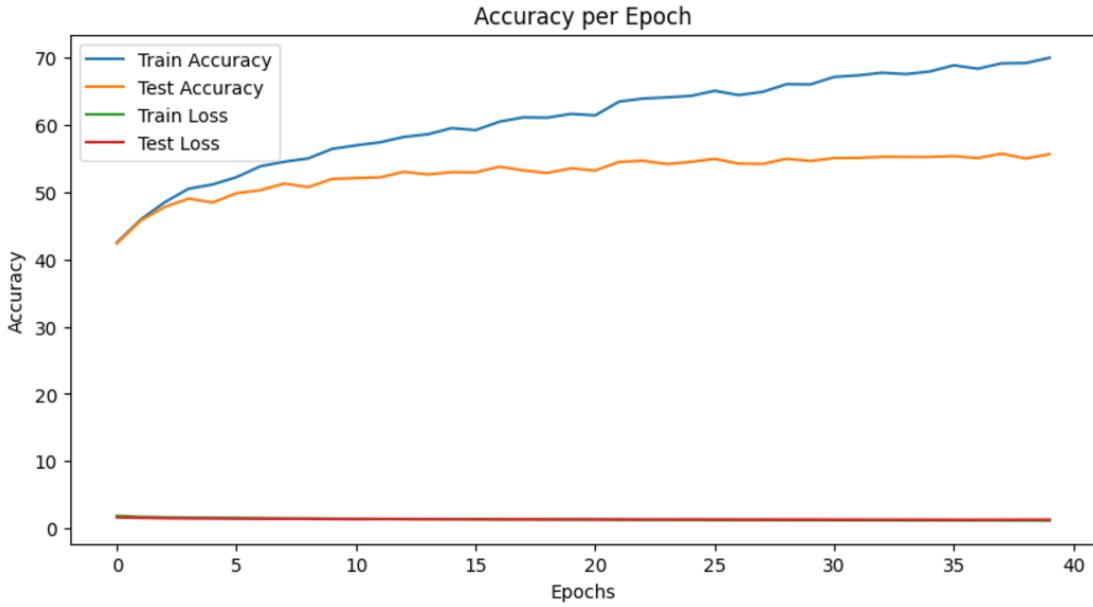
```

```

Train - Loss: 1.2146, Accuracy: 67.18%, Precision: 0.68, Recall: 0.67, F1: 0.67
Test - Loss: 1.2987, Accuracy: 55.11%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [32/40]
Train - Loss: 1.2109, Accuracy: 67.41%, Precision: 0.68, Recall: 0.67, F1: 0.67
Test - Loss: 1.2981, Accuracy: 55.12%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [33/40]
Train - Loss: 1.2052, Accuracy: 67.80%, Precision: 0.69, Recall: 0.68, F1: 0.68
Test - Loss: 1.2866, Accuracy: 55.29%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [34/40]
Train - Loss: 1.1949, Accuracy: 67.60%, Precision: 0.68, Recall: 0.68, F1: 0.67
Test - Loss: 1.2943, Accuracy: 55.27%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [35/40]
Train - Loss: 1.1914, Accuracy: 68.00%, Precision: 0.69, Recall: 0.68, F1: 0.68
Test - Loss: 1.2915, Accuracy: 55.26%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [36/40]
Train - Loss: 1.1879, Accuracy: 68.90%, Precision: 0.70, Recall: 0.69, F1: 0.69
Test - Loss: 1.2874, Accuracy: 55.39%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [37/40]
Train - Loss: 1.1782, Accuracy: 68.40%, Precision: 0.69, Recall: 0.68, F1: 0.69
Test - Loss: 1.2863, Accuracy: 55.10%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [38/40]
Train - Loss: 1.1777, Accuracy: 69.20%, Precision: 0.70, Recall: 0.69, F1: 0.69
Test - Loss: 1.2847, Accuracy: 55.77%, Precision: 0.57, Recall: 0.56, F1: 0.56
Epoch [39/40]
Train - Loss: 1.1654, Accuracy: 69.24%, Precision: 0.71, Recall: 0.69, F1: 0.69
Test - Loss: 1.2977, Accuracy: 55.03%, Precision: 0.56, Recall: 0.55, F1: 0.55
Epoch [40/40]
Train - Loss: 1.1584, Accuracy: 70.03%, Precision: 0.70, Recall: 0.70, F1: 0.70
Test - Loss: 1.2958, Accuracy: 55.69%, Precision: 0.56, Recall: 0.56, F1: 0.56

Final Evaluation:
Training Set - Accuracy: 70.03%, Precision: 0.70, Recall: 0.70, F1: 0.70
Test Set - Accuracy: 55.69%, Precision: 0.56, Recall: 0.56, F1: 0.56

```



Η συμπεριφορά των accuracies είναι παρόμοιες με του προηγούμενου πειράματος. Στα τελευταία 4 πειράματα έχουμε το test accuracy να μην μεταβάλλεται σημαντικά (διαφορές < 1%) ενώ από την άλλη το train accuracy να μειώνεται. Ίσως πρέπει να αλλάξουμε κάποιες παραμέτρους. Στο επόμενο πείραμα θα προσδώσουμε περισσότερο βάθος στο νευρωνικό προσθέτοντας ένα ακόμα layer αποσκοπώντας σε καλύτερη αναπαράσταση των δεδομένων της cifar-10 με τους νευρώνες να διαμορφώνονται ως εξής: 1024 στο πρώτο hidden

layer, 512 στο δεύτερο, 256 στο τρίτο και 128 στο τέταρτο. Επίσης θα μειώσουμε το dropout από 0.5 σε 0.4 και θα αυξήσουμε το weight decay από 0.00001 σε 0.0001 ώστε να μειώσουμε τη πιθανότητα overfitting. Τέλος, θα αυξήσουμε το batch size σε 80 το οποίο αποτελεί μια μέση τιμή των πειραμάτων μας και θα αυξήσουμε το learning rate σε 0.001 και τις εποχές από 40 σε 50 χρησιμοποιώντας πλέον και early stopping με patience=5 (θα σταματά την εκπαίδευση αν για 5 εποχές δεν έχουμε νέα καλύτερη τιμή για το test accuracy).

Η εκπαίδευση ολοκληρώθηκε σε 46 λεπτά και οι αποδόσεις του μοντέλου ήταν οι εξής:

```
Epoch [1/50]
Train - Loss: 1.8821, Accuracy: 41.01%, Precision: 0.40, Recall: 0.41, F1: 0.39
Test - Loss: 1.6977, Accuracy: 41.05%, Precision: 0.41, Recall: 0.41, F1: 0.39
Epoch [2/50]
Train - Loss: 1.7521, Accuracy: 42.40%, Precision: 0.45, Recall: 0.42, F1: 0.42
Test - Loss: 1.6437, Accuracy: 41.16%, Precision: 0.44, Recall: 0.41, F1: 0.40
Epoch [3/50]
Train - Loss: 1.6942, Accuracy: 46.24%, Precision: 0.48, Recall: 0.46, F1: 0.46
Test - Loss: 1.5912, Accuracy: 44.80%, Precision: 0.46, Recall: 0.45, F1: 0.44
Epoch [4/50]
Train - Loss: 1.6642, Accuracy: 47.03%, Precision: 0.48, Recall: 0.47, F1: 0.46
Test - Loss: 1.5665, Accuracy: 44.46%, Precision: 0.45, Recall: 0.44, F1: 0.44
Epoch [5/50]
Train - Loss: 1.6418, Accuracy: 47.81%, Precision: 0.48, Recall: 0.48, F1: 0.47
Test - Loss: 1.5366, Accuracy: 46.11%, Precision: 0.46, Recall: 0.46, F1: 0.46
Epoch [6/50]
Train - Loss: 1.6298, Accuracy: 48.31%, Precision: 0.50, Recall: 0.48, F1: 0.48
Test - Loss: 1.5303, Accuracy: 46.37%, Precision: 0.48, Recall: 0.46, F1: 0.46
Epoch [7/50]
Train - Loss: 1.6152, Accuracy: 48.63%, Precision: 0.48, Recall: 0.49, F1: 0.47
Test - Loss: 1.4905, Accuracy: 47.23%, Precision: 0.46, Recall: 0.47, F1: 0.45
Epoch [8/50]
Train - Loss: 1.6072, Accuracy: 48.13%, Precision: 0.49, Recall: 0.48, F1: 0.48
Test - Loss: 1.5222, Accuracy: 45.85%, Precision: 0.47, Recall: 0.46, F1: 0.45
Epoch [9/50]
Train - Loss: 1.5996, Accuracy: 49.37%, Precision: 0.52, Recall: 0.49, F1: 0.49
Test - Loss: 1.5034, Accuracy: 47.09%, Precision: 0.50, Recall: 0.47, F1: 0.47
Epoch [10/50]
Train - Loss: 1.5925, Accuracy: 50.08%, Precision: 0.50, Recall: 0.50, F1: 0.50
Test - Loss: 1.5123, Accuracy: 48.04%, Precision: 0.48, Recall: 0.48, F1: 0.47
Epoch [11/50]
Train - Loss: 1.5806, Accuracy: 49.96%, Precision: 0.51, Recall: 0.50, F1: 0.49
Test - Loss: 1.5064, Accuracy: 47.40%, Precision: 0.48, Recall: 0.47, F1: 0.47
Epoch [12/50]
Train - Loss: 1.5768, Accuracy: 49.72%, Precision: 0.50, Recall: 0.50, F1: 0.49
Test - Loss: 1.5113, Accuracy: 48.06%, Precision: 0.49, Recall: 0.48, F1: 0.47
Epoch [13/50]
Train - Loss: 1.5713, Accuracy: 48.60%, Precision: 0.50, Recall: 0.49, F1: 0.48
```

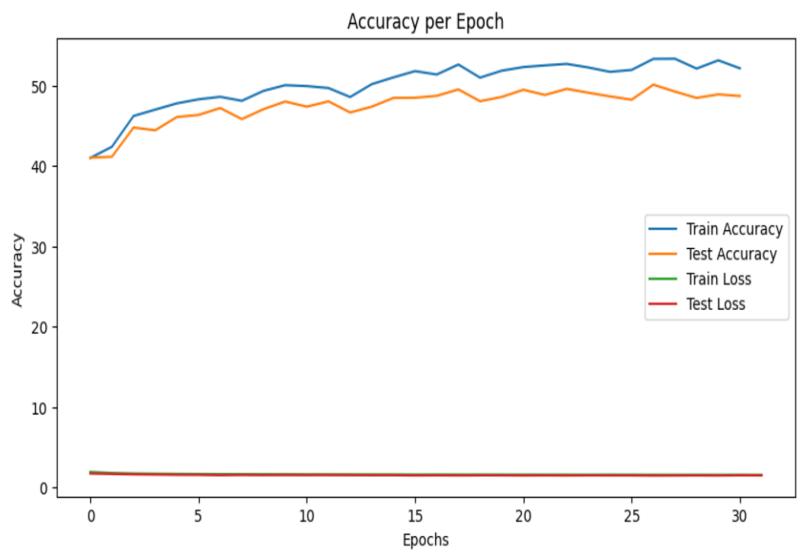
```
Epoch [13/50]
Train - Loss: 1.5713, Accuracy: 48.60%, Precision: 0.50, Recall: 0.49, F1: 0.48
Test - Loss: 1.5005, Accuracy: 46.68%, Precision: 0.48, Recall: 0.47, F1: 0.45
Epoch [14/50]
Train - Loss: 1.5737, Accuracy: 50.20%, Precision: 0.51, Recall: 0.50, F1: 0.49
Test - Loss: 1.4863, Accuracy: 47.40%, Precision: 0.48, Recall: 0.47, F1: 0.46
Epoch [15/50]
Train - Loss: 1.5682, Accuracy: 51.05%, Precision: 0.52, Recall: 0.51, F1: 0.50
Test - Loss: 1.4867, Accuracy: 48.49%, Precision: 0.50, Recall: 0.48, F1: 0.48
Epoch [16/50]
Train - Loss: 1.5601, Accuracy: 51.83%, Precision: 0.52, Recall: 0.52, F1: 0.51
Test - Loss: 1.4617, Accuracy: 48.51%, Precision: 0.48, Recall: 0.49, F1: 0.48
Epoch [17/50]
Train - Loss: 1.5629, Accuracy: 51.41%, Precision: 0.52, Recall: 0.51, F1: 0.51
Test - Loss: 1.4737, Accuracy: 48.75%, Precision: 0.49, Recall: 0.49, F1: 0.49
Epoch [18/50]
Train - Loss: 1.5600, Accuracy: 52.65%, Precision: 0.53, Recall: 0.53, F1: 0.53
Test - Loss: 1.4593, Accuracy: 49.55%, Precision: 0.50, Recall: 0.50, F1: 0.49
Epoch [19/50]
Train - Loss: 1.5556, Accuracy: 51.02%, Precision: 0.50, Recall: 0.51, F1: 0.48
Test - Loss: 1.4748, Accuracy: 48.08%, Precision: 0.47, Recall: 0.48, F1: 0.45
Epoch [20/50]
Train - Loss: 1.5523, Accuracy: 51.88%, Precision: 0.52, Recall: 0.52, F1: 0.51
Test - Loss: 1.4712, Accuracy: 48.61%, Precision: 0.49, Recall: 0.49, F1: 0.48
Epoch [21/50]
Train - Loss: 1.5514, Accuracy: 52.34%, Precision: 0.54, Recall: 0.52, F1: 0.52
Test - Loss: 1.4559, Accuracy: 49.51%, Precision: 0.51, Recall: 0.50, F1: 0.49
Epoch [22/50]
Train - Loss: 1.5460, Accuracy: 52.54%, Precision: 0.53, Recall: 0.53, F1: 0.52
Test - Loss: 1.4682, Accuracy: 48.86%, Precision: 0.49, Recall: 0.49, F1: 0.48
Epoch [23/50]
Train - Loss: 1.5455, Accuracy: 52.73%, Precision: 0.53, Recall: 0.53, F1: 0.52
```

```

Test - Loss: 1.4559, Accuracy: 49.51%, Precision: 0.51, Recall: 0.50, F1: 0.49
Epoch [22/50]
Train - Loss: 1.5460, Accuracy: 52.54%, Precision: 0.53, Recall: 0.53, F1: 0.52
Test - Loss: 1.4682, Accuracy: 48.86%, Precision: 0.49, Recall: 0.49, F1: 0.48
Epoch [23/50]
Train - Loss: 1.5455, Accuracy: 52.73%, Precision: 0.53, Recall: 0.53, F1: 0.52
Test - Loss: 1.4556, Accuracy: 49.62%, Precision: 0.50, Recall: 0.50, F1: 0.49
Epoch [24/50]
Train - Loss: 1.5371, Accuracy: 52.28%, Precision: 0.52, Recall: 0.52, F1: 0.52
Test - Loss: 1.4698, Accuracy: 49.14%, Precision: 0.49, Recall: 0.49, F1: 0.49
Epoch [25/50]
Train - Loss: 1.5456, Accuracy: 51.74%, Precision: 0.54, Recall: 0.52, F1: 0.51
Test - Loss: 1.4626, Accuracy: 48.68%, Precision: 0.50, Recall: 0.49, F1: 0.48
Epoch [26/50]
Train - Loss: 1.5414, Accuracy: 51.98%, Precision: 0.53, Recall: 0.52, F1: 0.51
Test - Loss: 1.4621, Accuracy: 48.27%, Precision: 0.50, Recall: 0.48, F1: 0.47
Epoch [27/50]
Train - Loss: 1.5328, Accuracy: 53.35%, Precision: 0.54, Recall: 0.53, F1: 0.53
Test - Loss: 1.4407, Accuracy: 50.14%, Precision: 0.50, Recall: 0.50, F1: 0.49
Epoch [28/50]
Train - Loss: 1.5330, Accuracy: 53.37%, Precision: 0.53, Recall: 0.53, F1: 0.53
Test - Loss: 1.4487, Accuracy: 49.28%, Precision: 0.50, Recall: 0.49, F1: 0.49
Epoch [29/50]
Train - Loss: 1.5324, Accuracy: 52.14%, Precision: 0.53, Recall: 0.52, F1: 0.51
Test - Loss: 1.4634, Accuracy: 48.49%, Precision: 0.49, Recall: 0.48, F1: 0.47
Epoch [30/50]
Train - Loss: 1.5319, Accuracy: 53.17%, Precision: 0.53, Recall: 0.53, F1: 0.52
Test - Loss: 1.4496, Accuracy: 48.93%, Precision: 0.49, Recall: 0.49, F1: 0.48
Epoch [31/50]
Train - Loss: 1.5309, Accuracy: 52.18%, Precision: 0.54, Recall: 0.52, F1: 0.52
Test - Loss: 1.4756, Accuracy: 48.73%, Precision: 0.50, Recall: 0.49, F1: 0.48
Early stopping at epoch 32

```

Final Evaluation:
 Training Set - Accuracy: 53.25%, Precision: 0.54, Recall: 0.53, F1: 0.53
 Test Set - Accuracy: 49.46%, Precision: 0.50, Recall: 0.49, F1: 0.49



Αρχικά πρέπει να επισημανθεί ότι λόγω λογικού λάθους που διορθώθηκε στη συνέχεια δεν εμφανίστηκαν τα στατιστικά της 32^η εποχής όπως έπρεπε και μετά να εκτελεστεί early stopping. Είναι άξιο προσοχής ότι το μοντέλο δυσκολεύεται ακόμα και στην εκπαίδευση, δηλαδή έχουμε σημάδια υποεκπαίδευσης. Αυτό ίσως να οφείλεται στην αλλαγή του learning rate από 0.0005 σε 0.001 ή στην αύξηση του weight decay από 0.000001 σε 0.0001. Επίσης, μπορεί για το συγκεκριμένο νευρωνικό δίκτυο αυτή η αύξηση του βάθους σε 3 hidden layers να είναι υπερβολική και μερικοί νευρώνες του τελευταίου layer να καταλήγουν να μη παίρνουν τιμή. Παρ'όλα αυτά θα το διατηρήσουμε και για το επόμενο πείραμα, στο οποίο θα έχουμε μείωση του learning rate από 0.001 σε 0.0004, αύξηση του batch size από 80 σε 90, αύξηση των εποχών σε 70 και του patience σε 8, καθώς έχουμε αρκετές εποχές και στοχεύουμε σε αργότερη και αποδοτικότερη εκπαίδευση λόγω του learning rate μας και της μείωσης του weight decay σε 0.000001.

Η εκπαίδευση επετεύχθη σε 63 λεπτά. Η απόδοση του μοντέλου φαίνεται παρακάτω:

Epoch [1/70]
Train - Loss: 1.8695, Accuracy: 43.41%, Precision: 0.43, Recall: 0.43, F1: 0.43
Test - Loss: 1.6184, Accuracy: 43.27%, Precision: 0.43, Recall: 0.43, F1: 0.43
Epoch [2/70]
Train - Loss: 1.6671, Accuracy: 47.58%, Precision: 0.47, Recall: 0.48, F1: 0.46
Test - Loss: 1.5104, Accuracy: 46.53%, Precision: 0.46, Recall: 0.47, F1: 0.45
Epoch [3/70]
Train - Loss: 1.5840, Accuracy: 50.07%, Precision: 0.51, Recall: 0.50, F1: 0.50
Test - Loss: 1.4768, Accuracy: 48.11%, Precision: 0.49, Recall: 0.48, F1: 0.47
Epoch [4/70]
Train - Loss: 1.5264, Accuracy: 52.87%, Precision: 0.53, Recall: 0.53, F1: 0.53
Test - Loss: 1.4411, Accuracy: 50.60%, Precision: 0.50, Recall: 0.51, F1: 0.50
Epoch [5/70]
Train - Loss: 1.4747, Accuracy: 53.81%, Precision: 0.54, Recall: 0.54, F1: 0.53
Test - Loss: 1.4052, Accuracy: 50.57%, Precision: 0.51, Recall: 0.51, F1: 0.51
Epoch [6/70]
Train - Loss: 1.4326, Accuracy: 56.26%, Precision: 0.56, Recall: 0.56, F1: 0.56
Test - Loss: 1.3570, Accuracy: 52.22%, Precision: 0.52, Recall: 0.52, F1: 0.51
Epoch [7/70]
Train - Loss: 1.3957, Accuracy: 56.38%, Precision: 0.57, Recall: 0.56, F1: 0.56
Test - Loss: 1.3802, Accuracy: 51.57%, Precision: 0.52, Recall: 0.52, F1: 0.51
Epoch [8/70]
Train - Loss: 1.3641, Accuracy: 58.88%, Precision: 0.59, Recall: 0.59, F1: 0.59
Test - Loss: 1.3403, Accuracy: 53.55%, Precision: 0.54, Recall: 0.54, F1: 0.53
Epoch [9/70]
Train - Loss: 1.3337, Accuracy: 59.87%, Precision: 0.60, Recall: 0.60, F1: 0.60
Test - Loss: 1.3181, Accuracy: 54.16%, Precision: 0.54, Recall: 0.54, F1: 0.54
Epoch [10/70]
Train - Loss: 1.3081, Accuracy: 61.22%, Precision: 0.61, Recall: 0.61, F1: 0.61
Test - Loss: 1.3086, Accuracy: 54.58%, Precision: 0.55, Recall: 0.55, F1: 0.54
Epoch [11/70]
Train - Loss: 1.2739, Accuracy: 62.24%, Precision: 0.62, Recall: 0.62, F1: 0.62
Test - Loss: 1.3059, Accuracy: 54.42%, Precision: 0.54, Recall: 0.54, F1: 0.54
Epoch [12/70]

Epoch [13/70]
Train - Loss: 1.2562, Accuracy: 63.51%, Precision: 0.64, Recall: 0.64, F1: 0.63
Test - Loss: 1.2955, Accuracy: 54.88%, Precision: 0.55, Recall: 0.55, F1: 0.54
Epoch [14/70]
Train - Loss: 1.2345, Accuracy: 64.35%, Precision: 0.65, Recall: 0.64, F1: 0.64
Test - Loss: 1.2810, Accuracy: 55.21%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [15/70]
Train - Loss: 1.2103, Accuracy: 64.79%, Precision: 0.65, Recall: 0.65, F1: 0.64
Test - Loss: 1.2798, Accuracy: 54.93%, Precision: 0.55, Recall: 0.55, F1: 0.54
Epoch [16/70]
Train - Loss: 1.1868, Accuracy: 66.40%, Precision: 0.66, Recall: 0.66, F1: 0.66
Test - Loss: 1.2700, Accuracy: 55.46%, Precision: 0.55, Recall: 0.55, F1: 0.55
Epoch [17/70]
Train - Loss: 1.1691, Accuracy: 67.20%, Precision: 0.68, Recall: 0.67, F1: 0.67
Test - Loss: 1.2670, Accuracy: 55.61%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [18/70]
Train - Loss: 1.1346, Accuracy: 68.46%, Precision: 0.68, Recall: 0.68, F1: 0.68
Test - Loss: 1.2796, Accuracy: 55.74%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [19/70]
Train - Loss: 1.1155, Accuracy: 69.44%, Precision: 0.69, Recall: 0.69, F1: 0.69
Test - Loss: 1.2666, Accuracy: 56.16%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [20/70]
Train - Loss: 1.0977, Accuracy: 70.59%, Precision: 0.71, Recall: 0.71, F1: 0.70
Test - Loss: 1.2627, Accuracy: 56.27%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [21/70]
Train - Loss: 1.0804, Accuracy: 71.44%, Precision: 0.72, Recall: 0.71, F1: 0.71
Test - Loss: 1.2554, Accuracy: 56.58%, Precision: 0.57, Recall: 0.57, F1: 0.56
Epoch [22/70]
Train - Loss: 1.0677, Accuracy: 71.92%, Precision: 0.72, Recall: 0.72, F1: 0.72
Test - Loss: 1.2613, Accuracy: 56.34%, Precision: 0.57, Recall: 0.56, F1: 0.56
Epoch [23/70]
Train - Loss: 1.0536, Accuracy: 72.59%, Precision: 0.73, Recall: 0.73, F1: 0.73
Test - Loss: 1.2592, Accuracy: 56.52%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [24/70]
Train - Loss: 1.0340, Accuracy: 73.59%, Precision: 0.74, Recall: 0.74, F1: 0.74
Test - Loss: 1.2594, Accuracy: 56.79%, Precision: 0.57, Recall: 0.57, F1: 0.57

Epoch [25/70]
Train - Loss: 1.0243, Accuracy: 73.77%, Precision: 0.74, Recall: 0.74, F1: 0.74
Test - Loss: 1.2602, Accuracy: 56.31%, Precision: 0.56, Recall: 0.56, F1: 0.56
Epoch [26/70]
Train - Loss: 1.0104, Accuracy: 74.51%, Precision: 0.75, Recall: 0.75, F1: 0.74
Test - Loss: 1.2691, Accuracy: 57.16%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [27/70]
Train - Loss: 1.0001, Accuracy: 75.01%, Precision: 0.75, Recall: 0.75, F1: 0.75
Test - Loss: 1.2625, Accuracy: 56.80%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [28/70]
Train - Loss: 0.9875, Accuracy: 75.07%, Precision: 0.75, Recall: 0.75, F1: 0.75
Test - Loss: 1.2669, Accuracy: 56.89%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [29/70]
Train - Loss: 0.9760, Accuracy: 75.76%, Precision: 0.76, Recall: 0.76, F1: 0.76
Test - Loss: 1.2708, Accuracy: 56.87%, Precision: 0.58, Recall: 0.57, F1: 0.57
Epoch [30/70]
Train - Loss: 0.9580, Accuracy: 76.19%, Precision: 0.76, Recall: 0.76, F1: 0.76
Test - Loss: 1.2681, Accuracy: 57.38%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [31/70]
Train - Loss: 0.9501, Accuracy: 77.36%, Precision: 0.77, Recall: 0.77, F1: 0.77
Test - Loss: 1.2565, Accuracy: 57.53%, Precision: 0.58, Recall: 0.58, F1: 0.57
Epoch [32/70]
Train - Loss: 0.9303, Accuracy: 77.75%, Precision: 0.78, Recall: 0.78, F1: 0.78
Test - Loss: 1.2619, Accuracy: 57.17%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [33/70]
Train - Loss: 0.9302, Accuracy: 78.81%, Precision: 0.79, Recall: 0.79, F1: 0.79
Test - Loss: 1.2700, Accuracy: 57.59%, Precision: 0.58, Recall: 0.58, F1: 0.58
Epoch [34/70]
Train - Loss: 0.9174, Accuracy: 78.10%, Precision: 0.78, Recall: 0.78, F1: 0.78
Test - Loss: 1.2728, Accuracy: 57.39%, Precision: 0.58, Recall: 0.57, F1: 0.57
Epoch [35/70]
Train - Loss: 0.9140, Accuracy: 79.39%, Precision: 0.80, Recall: 0.79, F1: 0.79
Test - Loss: 1.2651, Accuracy: 57.61%, Precision: 0.58, Recall: 0.58, F1: 0.57
Epoch [36/70]
Train - Loss: 0.8939, Accuracy: 79.69%, Precision: 0.80, Recall: 0.80, F1: 0.80
Test - Loss: 1.2747, Accuracy: 56.83%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [37/70]
Train - Loss: 0.8846, Accuracy: 80.02%, Precision: 0.80, Recall: 0.80, F1: 0.80
Test - Loss: 1.2801, Accuracy: 57.54%, Precision: 0.58, Recall: 0.58, F1: 0.58

```
Test - Loss: 1.2700, Accuracy: 57.55%, Precision: 0.58, Recall: 0.58, F1: 0.58
Epoch [40/70]
Train - Loss: 0.8626, Accuracy: 81.57%, Precision: 0.82, Recall: 0.82, F1: 0.82
Test - Loss: 1.2833, Accuracy: 57.93%, Precision: 0.58, Recall: 0.58, F1: 0.58
Epoch [41/70]
Train - Loss: 0.8541, Accuracy: 81.79%, Precision: 0.82, Recall: 0.82, F1: 0.82
Test - Loss: 1.2796, Accuracy: 57.95%, Precision: 0.59, Recall: 0.58, F1: 0.58
Epoch [42/70]
Train - Loss: 0.8424, Accuracy: 82.11%, Precision: 0.83, Recall: 0.82, F1: 0.82
Test - Loss: 1.2961, Accuracy: 57.17%, Precision: 0.58, Recall: 0.57, F1: 0.57
Epoch [43/70]
Train - Loss: 0.8277, Accuracy: 82.29%, Precision: 0.83, Recall: 0.82, F1: 0.82
Test - Loss: 1.2900, Accuracy: 57.58%, Precision: 0.58, Recall: 0.58, F1: 0.57
Epoch [44/70]
Train - Loss: 0.8242, Accuracy: 82.62%, Precision: 0.83, Recall: 0.83, F1: 0.83
Test - Loss: 1.2885, Accuracy: 57.68%, Precision: 0.58, Recall: 0.58, F1: 0.58
Epoch [45/70]
Train - Loss: 0.8207, Accuracy: 83.19%, Precision: 0.83, Recall: 0.83, F1: 0.83
Test - Loss: 1.3011, Accuracy: 57.63%, Precision: 0.58, Recall: 0.58, F1: 0.57
Epoch [46/70]
Train - Loss: 0.8099, Accuracy: 82.99%, Precision: 0.83, Recall: 0.83, F1: 0.83
Test - Loss: 1.3158, Accuracy: 57.16%, Precision: 0.57, Recall: 0.57, F1: 0.57
Epoch [47/70]
Train - Loss: 0.8057, Accuracy: 83.85%, Precision: 0.84, Recall: 0.84, F1: 0.84
Test - Loss: 1.3008, Accuracy: 57.48%, Precision: 0.58, Recall: 0.57, F1: 0.57
Epoch [48/70]
Train - Loss: 0.7939, Accuracy: 84.04%, Precision: 0.84, Recall: 0.84, F1: 0.84
Test - Loss: 1.3350, Accuracy: 57.53%, Precision: 0.58, Recall: 0.58, F1: 0.57
Epoch [49/70]
Train - Loss: 0.7911, Accuracy: 84.43%, Precision: 0.85, Recall: 0.84, F1: 0.85
Test - Loss: 1.3200, Accuracy: 57.14%, Precision: 0.58, Recall: 0.57, F1: 0.57
Early stopping at epoch 49

Final Evaluation:
Training Set - Accuracy: 84.43%, Precision: 0.85, Recall: 0.84, F1: 0.85
Test Set      - Accuracy: 57.14%, Precision: 0.58, Recall: 0.57, F1: 0.57
```

To early stopping που

προστέθηκε στο κώδικα συνέβαλε στο να μην εκτελεστεί το πρόγραμμα για ακόμη 21 εποχές. Όπως φαίνεται, το train loss μειώνεται σταδιακά καθώς και το train accuracy αυξάνεται όσο συνεχίζει η εκπαίδευση. Το test accuracy έχει ένα ταβάνι στα 57.95% το οποίο αποτελεί μία καλή απόδοση για MLP στη βάση δεδομένων cifar-10. Αυτό για το οποίο πρέπει να προβληματιστούμε είναι η πορεία του test loss. Στις πρώτες εποχές, το μοντέλο μαθαίνει τις βασικές δομές των δεδομένων και αρχίζει να γενικεύει καλά και έτσι το test loss μειώνεται. Από κάποιο σημείο και μετά όμως, ενώ το train loss

συνεχίζει σταδιακά να μειώνεται, το test loss αρχίζει να αυξάνει. Αυτό μπορεί να σημαίνει ότι το μοντέλο υπερπροσαρμόζεται (overfitting), δηλαδή μαθαίνει χαρακτηριστικά που είναι συγκεκριμένα για τα δεδομένα εκπαίδευσης και δεν γενικεύουν καλά στο test set.