```
import torch
import torch.nn as nn
import torch.optim as optim
import torchvision.transforms as transforms
import torchvision.datasets as datasets
from torch.utils.data import DataLoader, random_split
from timm import create_model
import os
import matplotlib.pyplot as plt
from sklearn.metrics import accuracy_score, roc_auc_score
from PIL import Image
```

```
import torch
device = torch.device("cpu")
print("Using CPU for training. It will be slow but will work.")

# Paths
DATA_DIR = "/content/drive/MyDrive/CAPSTONE/Dataset_patches_small" #
Folder containing 'IDC' and 'normal' subfolders
BATCH_SIZE = 32
EPOCHS = 10
LR = 0.001
IMG_SIZE = 256
NUM_CLASSES = 2 # IDC vs Normal
```

→ Using CPU for training. It will be slow but will work.

```
train_transform = transforms.Compose([
    transforms.Resize((IMG_SIZE, IMG_SIZE)),
    transforms.RandomHorizontalFlip(),
    transforms.RandomRotation(15),
    transforms.ColorJitter(brightness=0.2, contrast=0.2),
    transforms.ToTensor(),
    transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224,
0.225])
])

val_transform = transforms.Compose([
    transforms.Resize((IMG_SIZE, IMG_SIZE)),
    transforms.ToTensor(),
```

```
transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224,
0.225])
full dataset = datasets.ImageFolder(root=DATA DIR,
transform=train transform)
train_size = int(0.8 * len(full dataset))
val size = int(0.1 * len(full dataset))
test size = len(full dataset) - train size - val size
train_dataset,    val_dataset,    test_dataset = random_split(full_dataset,
[train size, val size, test size])
train loader = DataLoader(train dataset, batch size=BATCH SIZE,
shuffle=True, num workers=4)
val loader = DataLoader(val dataset, batch size=BATCH SIZE,
shuffle=False, num workers=4)
test loader = DataLoader(test dataset, batch size=BATCH SIZE,
shuffle=False, num workers=4)
{len(test dataset)}")
   Train: 960, Val: 120, Test: 120
/usr/local/lib/python3.11/dist-packages/torch/utils/data/dataloader.py:624: UserWarning: This DataLoader will create 4 worker processes in total. Our suggestable of the control of the 
model = create model("efficientnet b0", pretrained=True,
num classes=NUM CLASSES)
model.to(device)
criterion = nn.CrossEntropyLoss()
optimizer = optim.Adam(model.parameters(), lr=LR)
 /usr/local/lib/python3.11/dist-packages/huggingface_hub/utils/_auth.py:94: UserWarning:
The secret 'HF_TOKEN' does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab (<a href="https://huggingface.co/settings/tokens">https://huggingface.co/settings/tokens</a>), set it as secret in your Google Col You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access public models or datasets.
warnings.warn(
```

21.4M/21.4M [00:00<00:00, 49.1MB/s]

ndel safetensors: 100%

```
import torch
import torch.nn as nn
import torch.optim as optim
from torch.utils.data import DataLoader, random split
import torchvision.transforms as transforms
import torchvision.datasets as datasets
from timm import create model
import os
device = torch.device("cuda" if torch.cuda.is available() else "cpu")
print("Using", device)
DATA DIR = "/content/drive/MyDrive/CAPSTONE/Dataset patches small"
BATCH SIZE = 8
IMG SIZE = 224
NUM CLASSES = 2
EPOCHS = 15
LR = 0.0005
train transform = transforms.Compose([
  transforms.Resize((IMG SIZE, IMG SIZE)),
  transforms.RandomHorizontalFlip(),
  transforms.RandomRotation(15),
  transforms.ColorJitter(0.2, 0.2),
  transforms.ToTensor(),
  transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224,
0.2251)
val transform = transforms.Compose([
   transforms.Resize((IMG SIZE, IMG SIZE)),
  transforms.ToTensor(),
])
full dataset = datasets.ImageFolder(DATA DIR,
transform=train transform)
train size = int(0.8 * len(full dataset))
val_size = int(0.1 * len(full_dataset))
```

```
test_size = len(full_dataset) - train_size - val_size
train_ds, val_ds, test_ds = random_split(full_dataset, [train_size,
val size, test size])
train loader = DataLoader(train ds, batch size=BATCH SIZE,
shuffle=True, num workers=2)
val_loader = DataLoader(val_ds, batch_size=BATCH_SIZE, shuffle=False,
num workers=2)
test_loader = DataLoader(test_ds, batch_size=BATCH_SIZE, shuffle=False,
num workers=2)
model = create model("efficientnet b0", pretrained=True,
num classes=NUM CLASSES)
model = model.to(device)
# Loss and optimizer
criterion = nn.CrossEntropyLoss()
optimizer = optim.Adam(model.parameters(), lr=LR)
def train(model, train loader, val loader, optimizer, criterion,
epochs):
  best acc = 0.0
  for epoch in range (epochs):
      model.train()
      running_loss, correct, total = 0.0, 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()
           preds = torch.argmax(outputs, dim=1)
```

```
correct += (preds == labels).sum().item()
           total += labels.size(0)
       train acc = correct / total
       print(f"Epoch [{epoch+1}/{epochs}] - Loss:
{running loss/len(train loader):.4f}, Train Acc: {train acc:.4f}")
      model.eval()
           for images, labels in val loader:
               images, labels = images.to(device), labels.to(device)
               outputs = model(images)
              preds = torch.argmax(outputs, dim=1)
               val correct += (preds == labels).sum().item()
              val total += labels.size(0)
      print(f"Validation Acc: {val acc:.4f}")
          best acc = val acc
           torch.save(model.state dict(),
"/content/drive/MyDrive/CAPSTONE/best model.pth")
train(model, train loader, val loader, optimizer, criterion, EPOCHS)
```

```
Epoch [1/15] - Loss: 1.9253, Train Acc: 0.7406 Validation Acc: 0.8417
    Best model saved!
    Epoch [2/15] - Loss: 1.0016, Train Acc: 0.8333
    Validation Acc: 0.9167
    Best model saved!
    Epoch [3/15] - Loss: 0.9667, Train Acc: 0.8271
    Validation Acc: 0.8583
    Epoch [4/15] - Loss: 0.5746, Train Acc: 0.8604
    Validation Acc: 0.8917
    Epoch [5/15] - Loss: 0.4055, Train Acc: 0.8781
    Validation Acc: 0.8667
    Epoch [6/15] - Loss: 0.2747, Train Acc: 0.9031
    Validation Acc: 0.9083
    Epoch [7/15] - Loss: 0.2883, Train Acc: 0.9000
    Validation Acc: 0.9333
    Best model saved!
    Epoch [8/15] - Loss: 0.2051, Train Acc: 0.9323
    Validation Acc: 0.9250
    Epoch [9/15] - Loss: 0.1522, Train Acc: 0.9458
    Validation Acc: 0.9083
    Epoch [10/15] - Loss: 0.1591, Train Acc: 0.9385
    Validation Acc: 0.8917
    Epoch [11/15] - Loss: 0.1405, Train Acc: 0.9563
    Validation Acc: 0.9333
    Epoch [12/15] - Loss: 0.1997, Train Acc: 0.9479
    Validation Acc: 0.9250
    Epoch [13/15] - Loss: 0.1725, Train Acc: 0.9417
    Validation Acc: 0.9167
    Epoch [14/15] - Loss: 0.1351, Train Acc: 0.9552
    Validation Acc: 0.9250
    Epoch [15/15] - Loss: 0.1710, Train Acc: 0.9458
    Validation Acc: 0.8750
```

```
import matplotlib.pyplot as plt

def train(model, train_loader, val_loader, optimizer, criterion, epochs):
    best_acc = 0.0

# Track metrics
    train_acc_list, val_acc_list = [], []
    train_loss_list, val_loss_list = [], []

for epoch in range(epochs):
    model.train()
    running_loss, correct, total = 0.0, 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()
          preds = torch.argmax(outputs, dim=1)
          correct += (preds == labels).sum().item()
          total += labels.size(0)
      train_loss = running_loss / len(train_loader)
      train acc list.append(train acc)
      train loss list.append(train loss)
      model.eval()
      with torch.no grad():
          for images, labels in val_loader:
              images, labels = images.to(device), labels.to(device)
              outputs = model(images)
              loss = criterion(outputs, labels)
              preds = torch.argmax(outputs, dim=1)
              val loss += loss.item()
              val correct += (preds == labels).sum().item()
              val total += labels.size(0)
      val loss /= len(val loader)
      val acc list.append(val acc)
      val loss list.append(val loss)
      print(f"Epoch [{epoch+1}/{epochs}] - Train Loss:
{train loss:.4f}, Train Acc: {train acc:.4f} | Val Loss:
{val loss:.4f}, Val Acc: {val acc:.4f}")
      if val acc > best acc:
          torch.save(model.state dict(),
          print("V Best model saved!")
```

```
os.makedirs("/content/drive/MyDrive/CAPSTONE/metrics",
exist ok=True)
  plt.figure(figsize=(10, 5))
  plt.plot(train acc list, label="Train Accuracy")
  plt.plot(val acc list, label="Validation Accuracy")
  plt.xlabel("Epoch")
  plt.ylabel("Accuracy")
  plt.title("Training vs Validation Accuracy")
  plt.legend()
  plt.grid()
plt.savefig("/content/drive/MyDrive/CAPSTONE/metrics/accuracy curve.png
  plt.show()
  plt.figure(figsize=(10, 5))
  plt.plot(train loss list, label="Train Loss")
  plt.plot(val loss list, label="Validation Loss")
  plt.xlabel("Epoch")
  plt.ylabel("Loss")
  plt.title("Training vs Validation Loss")
  plt.legend()
  plt.grid()
plt.savefig("/content/drive/MyDrive/CAPSTONE/metrics/loss curve.png")
  plt.show()
import torch
import torch.nn as nn
import torch.optim as optim
from torch.utils.data import DataLoader, random split
import torchvision.transforms as transforms
import torchvision.datasets as datasets
from timm import create model
import os
import matplotlib.pyplot as plt
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
```

```
print("Using", device)
DATA DIR = "/content/drive/MyDrive/CAPSTONE/Dataset patches small"
SAVE DIR = "/content/drive/MyDrive/CAPSTONE"
METRIC DIR = os.path.join(SAVE DIR, "metrics")
BATCH SIZE = 8
IMG SIZE = 224
NUM CLASSES = 2
EPOCHS = 30
LR = 0.0005
os.makedirs(METRIC DIR, exist ok=True)
train transform = transforms.Compose([
  transforms.Resize((IMG SIZE, IMG SIZE)),
  transforms.RandomHorizontalFlip(),
  transforms.RandomRotation(15),
  transforms.ColorJitter(0.2, 0.2),
  transforms.ToTensor(),
  transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224,
0.225])
val transform = transforms.Compose([
  transforms.Resize((IMG SIZE, IMG SIZE)),
   transforms.ToTensor(),
   transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224,
0.225])
1)
full_dataset = datasets.ImageFolder(DATA_DIR,
transform=train transform)
train size = int(0.8 * len(full dataset))
val size = int(0.1 * len(full dataset))
test size = len(full dataset) - train size - val size
train_ds, val_ds, test_ds = random split(full dataset, [train size,
val size, test size])
```

```
train loader = DataLoader(train ds, batch size=BATCH SIZE,
shuffle=True, num workers=2)
val loader = DataLoader(val ds, batch size=BATCH SIZE, shuffle=False,
num workers=2)
test loader = DataLoader(test ds, batch size=BATCH SIZE, shuffle=False,
num workers=2)
model = create model("efficientnet b0", pretrained=True,
num classes=NUM CLASSES)
model = model.to(device)
criterion = nn.CrossEntropyLoss()
optimizer = optim.Adam(model.parameters(), lr=LR)
train accuracies = []
val accuracies = []
train losses = []
val losses = []
def train(model, train loader, val loader, optimizer, criterion,
epochs):
  for epoch in range (epochs):
      model.train()
       running loss, correct, total = 0.0, 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()
           preds = torch.argmax(outputs, dim=1)
           correct += (preds == labels).sum().item()
           total += labels.size(0)
```

```
train acc = correct / total
       train loss = running loss / len(train loader)
       train accuracies.append(train acc)
       train losses.append(train loss)
      print(f"Epoch [{epoch+1}/{epochs}] - Loss: {train loss:.4f},
Train Acc: {train acc:.4f}")
      model.eval()
      val correct, val total = 0, 0
      with torch.no grad():
           for images, labels in val loader:
               images, labels = images.to(device), labels.to(device)
               outputs = model(images)
              loss = criterion(outputs, labels)
              preds = torch.argmax(outputs, dim=1)
              val correct += (preds == labels).sum().item()
              val running loss += loss.item()
      val loss = val running loss / len(val loader)
       val accuracies.append(val acc)
       val losses.append(val loss)
       print(f"Validation Acc: {val acc:.4f}")
       if val acc > best acc:
           torch.save(model.state_dict(), os.path.join(SAVE_DIR,
"best model.pth"))
           print("V Best model saved!")
  plt.figure(figsize=(10, 5))
  plt.plot(train accuracies, label='Train Accuracy')
  plt.plot(val accuracies, label='Validation Accuracy')
  plt.xlabel('Epoch')
  plt.ylabel('Accuracy')
  plt.title('Accuracy Curve')
  plt.legend()
  plt.grid(True)
```

```
plt.savefig(os.path.join(METRIC_DIR, "accuracy_curve.png"))
     plt.close()
     plt.figure(figsize=(10, 5))
     plt.plot(val losses, label='Validation Loss')
     plt.xlabel('Epoch')
     plt.ylabel('Loss')
     plt.title('Loss Curve')
     plt.legend()
     plt.grid(True)
     plt.savefig(os.path.join(METRIC DIR, "loss curve.png"))
     plt.close()
train(model, train loader, val loader, optimizer, criterion, EPOCHS)
 model.safetensors: 100%
                                                               21.4M/21.4M [00:00<00:00, 80.2MB/s]
      Epoch [1/30] - Loss: 1.7011, Train Acc: 0.7635
Validation Acc: 0.7500

▼ Best model saved!
Epoch [2/30] - Loss: 1.2627, Train Acc: 0.8052

      Validation Acc: 0.8583

▼ Best model saved!
Epoch [3/30] - Loss: 0.6943, Train Acc: 0.8427

      Validation Acc: 0.7833
Epoch [4/30] − Loss: 0.9322, Train Acc: 0.8167
Validation Acc: 0.9000
✓ Best model saved!
Epoch [5/30] − Loss: 0.3769, Train Acc: 0.8927
      ✓ Best model saved!

Epoch [6/30] - Loss: 0.4407, Train Acc: 0.8771
      Validation Acc: 0.8750
Epoch [7/30] - Loss: 0.4176, Train Acc: 0.8833
      Validation Acc: 0.8833
Epoch [8/30] - Loss: 0.2931, Train Acc: 0.8969
Validation Acc: 0.9250
Epoch [9/30] - Loss: 0.2026, Train Acc: 0.9302
Validation Acc: 0.9250
```

Validation Acc: 0.9250
Epoch [10/30] - Loss: 0.2341, Train Acc: 0.9250
Validation Acc: 0.8833
Epoch [11/30] - Loss: 0.2040, Train Acc: 0.9323
Validation Acc: 0.9333
Fooch [12/30] - Loss: 0.1199, Train Acc: 0.9583
Validation Acc: 0.9333
Fooch [12/30] - Loss: 0.1670, Train Acc: 0.9333

Epoch [13/30] - Loss: 0.1670, Train Acc: 0.9323 Validation Acc: 0.9250 Epoch [14/30] - Loss: 0.2044, Train Acc: 0.9323 Validation Acc: 0.9167

```
Validation Acc: 0.9167
    Epoch [15/30] - Loss: 0.1529, Train Acc: 0.9500
→ Validation Acc: 0.8500
Epoch [16/30] - Loss: 0.1459, Train Acc: 0.9563
    Validation Acc: 0.9167
    Epoch [17/30] - Loss: 0.1428, Train Acc: 0.9542
    Validation Acc: 0.9167
    Epoch [18/30] - Loss: 0.0944, Train Acc: 0.9688
    Validation Acc: 0.9083
    Epoch [19/30] - Loss: 0.0836, Train Acc: 0.9708
    Validation Acc: 0.9333
    Epoch [20/30] - Loss: 0.1046, Train Acc: 0.9708
    Validation Acc: 0.9167
    Epoch [21/30] - Loss: 0.1149, Train Acc: 0.9625
    Validation Acc: 0.9250
    Epoch [22/30] - Loss: 0.0557, Train Acc: 0.9792
    Validation Acc: 0.8833
    Epoch [23/30] - Loss: 0.1177, Train Acc: 0.9635
    Validation Acc: 0.9167
    Epoch [24/30] - Loss: 0.1057, Train Acc: 0.9563
    Validation Acc: 0.9417

▼ Best model saved!

    Epoch [25/30] - Loss: 0.1604, Train Acc: 0.9458
    Validation Acc: 0.9250
    Epoch [26/30] - Loss: 0.1035, Train Acc: 0.9667
    Validation Acc: 0.8833
    Epoch [27/30] - Loss: 0.1106, Train Acc: 0.9688
    Validation Acc: 0.9167
    Epoch [28/30] - Loss: 0.0604, Train Acc: 0.9771
    Validation Acc: 0.8833
    Epoch [29/30] - Loss: 0.1915, Train Acc: 0.9375
    Validation Acc: 0.9167
    Epoch [30/30] - Loss: 0.0996, Train Acc: 0.9625
    Validation Acc: 0.8583
```

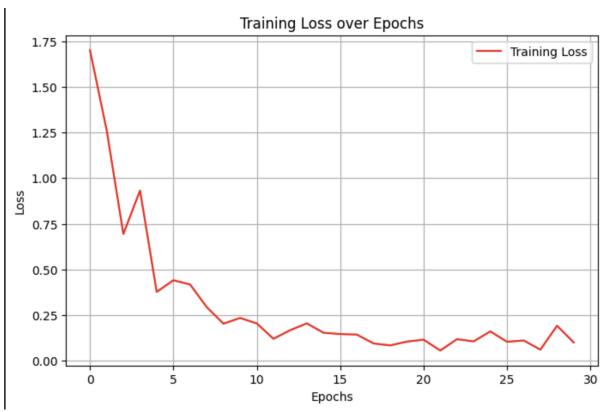
```
import matplotlib.pyplot as plt
import os

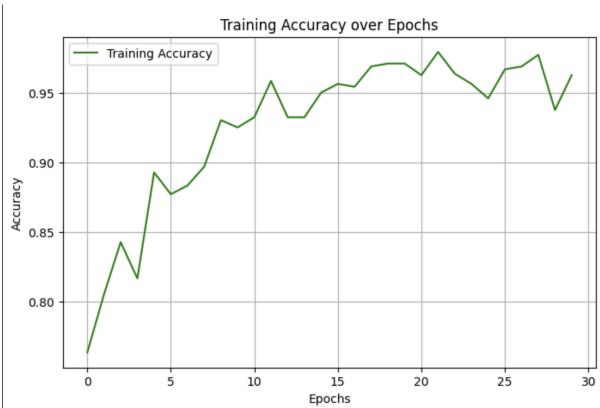
metrics_dir = "/content/drive/MyDrive/CAPSTONE/metrics"
os.makedirs(metrics_dir, exist_ok=True)

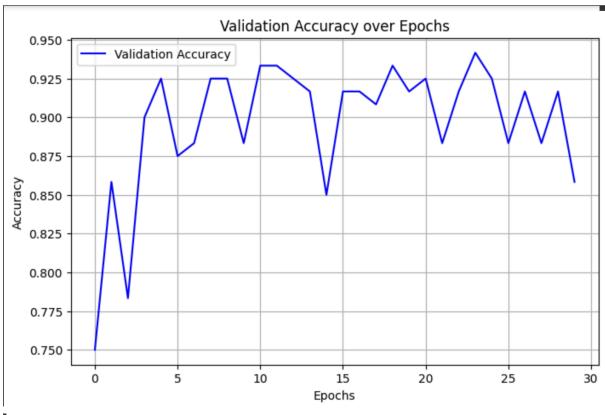
plt.figure(figsize=(8, 5))
plt.plot(train_losses, label='Training Loss', color='red')
plt.title("Training Loss over Epochs")
plt.xlabel("Epochs")
plt.ylabel("Loss")
plt.legend()
plt.grid(True)
plt.savefig(os.path.join(metrics_dir, "training_loss.png"))
plt.show()

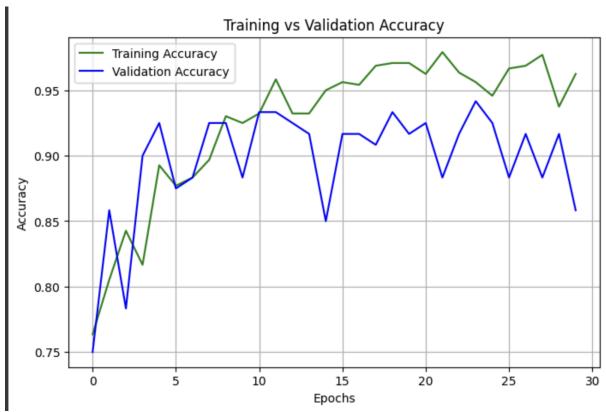
plt.figure(figsize=(8, 5))
plt.plot(train_accuracies, label='Training Accuracy', color='green')
plt.title("Training Accuracy over Epochs")
```

```
plt.xlabel("Epochs")
plt.ylabel("Accuracy")
plt.legend()
plt.grid(True)
plt.savefig(os.path.join(metrics dir, "training accuracy.png"))
plt.show()
plt.figure(figsize=(8, 5))
plt.plot(val accuracies, label='Validation Accuracy', color='blue')
plt.title("Validation Accuracy over Epochs")
plt.xlabel("Epochs")
plt.ylabel("Accuracy")
plt.legend()
plt.grid(True)
plt.savefig(os.path.join(metrics_dir, "validation_accuracy.png"))
plt.show()
plt.figure(figsize=(8, 5))
plt.plot(train accuracies, label='Training Accuracy', color='green')
plt.plot(val accuracies, label='Validation Accuracy', color='blue')
plt.title("Training vs Validation Accuracy")
plt.xlabel("Epochs")
plt.ylabel("Accuracy")
plt.legend()
plt.grid(True)
plt.savefig(os.path.join(metrics_dir,
"train_val_accuracy_comparison.png"))
plt.show()
```



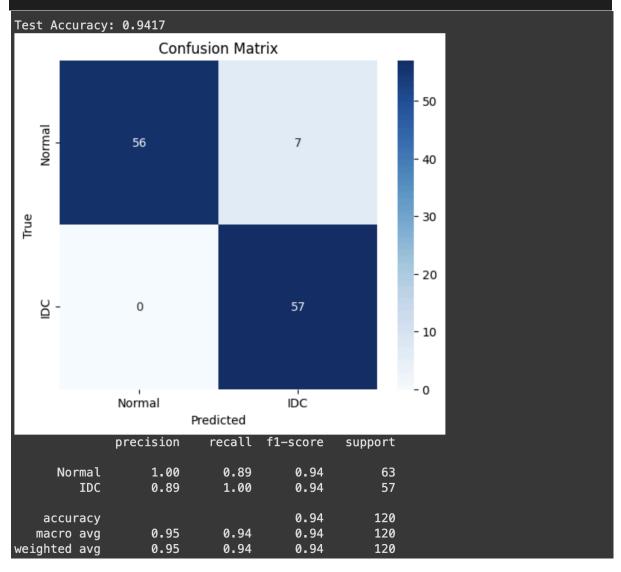






```
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics import confusion matrix, classification report,
accuracy score
import numpy as np
import os
os.makedirs("/content/drive/MyDrive/CAPSTONE/metrics", exist ok=True)
model.load state dict(torch.load("/content/drive/MyDrive/CAPSTONE/best
model.pth"))
model.eval()
all preds = []
all labels = []
with torch.no grad():
   for images, labels in test loader:
       images, labels = images.to(device), labels.to(device)
       outputs = model(images)
       preds = torch.argmax(outputs, dim=1)
       all preds.extend(preds.cpu().numpy())
       all labels.extend(labels.cpu().numpy())
acc = accuracy_score(all_labels, all_preds)
print(f"Test Accuracy: {acc:.4f}")
cm = confusion matrix(all labels, all preds)
plt.figure(figsize=(6, 5))
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues',
xticklabels=['Normal', 'IDC'], yticklabels=['Normal', 'IDC'])
plt.xlabel("Predicted")
plt.ylabel("True")
plt.title("Confusion Matrix")
plt.savefig("/content/drive/MyDrive/CAPSTONE/metrics/confusion matrix.p
ng")
plt.show()
```

```
report = classification_report(all_labels, all_preds,
target_names=['Normal', 'IDC'])
print(report)
with
open("/content/drive/MyDrive/CAPSTONE/metrics/classification_report.txt
", "w") as f:
    f.write(report)
```



```
import matplotlib.pyplot as plt
import cv2
import os

gradcam_dir = "/content/drive/MyDrive/CAPSTONE/gradcam"
```

```
gradcam_images = [f for f in os.listdir(gradcam_dir) if
f.endswith('.png')]

for img_name in gradcam_images:
   img_path = os.path.join(gradcam_dir, img_name)

img = cv2.imread(img_path)
   img_rgb = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)

plt.figure(figsize=(8, 6))
   plt.imshow(img_rgb)
   plt.axis('off')
   plt.title(f"Grad-CAM for {img_name}")
   plt.show()
```

