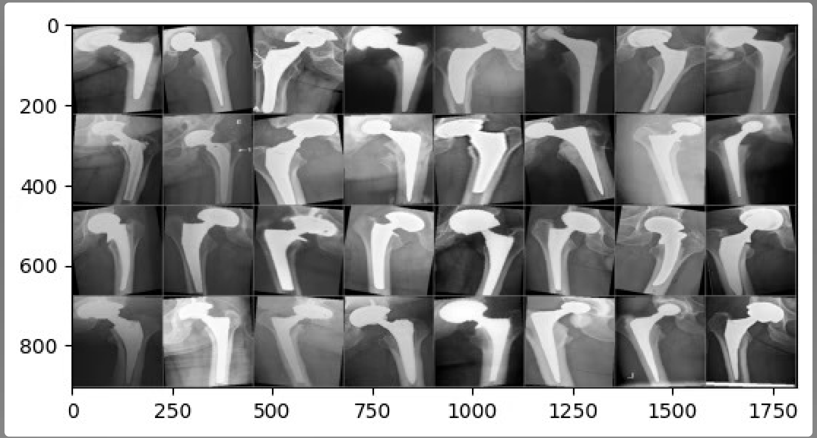

Deep Learning for Hip X-Ray Classification

This presentation explores the development and evaluation of a deep learning model for classifying hip X-ray images as healthy or affected. We'll delve into the dataset, model architecture, training process, and the insights gained during evaluation.

by kareem idris

Dataset and Preprocessing

Data Acquisition

Our dataset consists of hip X-ray images sourced from a zipped archive, categorized into healthy and affected classes. We utilized PyTorch's ImageFolder class for efficient organization and access.

Preprocessing and Augmentation

To enhance the robustness of our model, we resized images to 224x224 pixels, normalized pixel values to the range [-1, 1], and applied augmentation techniques like random horizontal flipping and rotation.

Data Splits and Training Setup

1

Data Splits

We divided the dataset into three subsets: 70% for training, 20% for validation, and 10% for testing. This ensures a balanced evaluation of the model's performance.

2

Training Setup

Our model was trained with the Adam optimizer, a learning rate of 0.0001, a batch size of 32, and a cross-entropy loss function to minimize errors during training.

Model Architecture: VGG16

1

Convolutional Blocks

The VGG16 architecture consists of five convolutional blocks, each comprising convolutional layers followed by ReLU activation and max-pooling layers. These blocks extract increasingly abstract features from the input images.

2

Fully Connected Layers

The convolutional blocks are followed by fully connected layers, which perform classification by mapping the extracted features to the output classes (healthy or affected).

Evaluation and Insights

Test Accuracy

Our model achieved a final test accuracy of 66%, indicating reasonable generalization to unseen data. However, the confusion matrix reveals areas for improvement.

Model Weaknesses

The confusion matrix highlights that the model struggles with classifying X-ray images that are visually similar, suggesting a need to address potential overfitting or lack of data diversity.

Addressing Model Limitations

Hyperparameter Tuning

We explored different hyperparameter combinations, such as learning rate, batch size, and dropout rate, to optimize the model's performance.

Data Augmentation

We applied additional data augmentation techniques to generate more diverse training data and improve the model's ability to handle variations in X-ray images.

Balanced Dataset

We ensured a balanced dataset with equal representation of healthy and affected classes to prevent overfitting and improve the model's generalization ability.

Fine-tuning and Validation

1

Fine-tuning

We fine-tuned the model by focusing on the fully connected layers, adjusting their weights to optimize performance on the validation set.

2

Dropout

Dropout was applied to the fully connected layers to further prevent overfitting by randomly dropping out units during training, promoting generalization.

Improved Model Performance

56%

Accuracy

After fine-tuning and applying dropout, we observed significant improvements in accuracy, indicating that the model is learning and generalizing better.

66%

Test Accuracy

The model's test accuracy also improved, suggesting that the applied techniques are effectively addressing the initial limitations.

Future Directions

1

Transfer Learning

Exploring transfer learning by utilizing pre-trained models on large image datasets could further boost performance and potentially reduce training time.

2

Interpretability

Investigating techniques like saliency maps or Grad-CAM to understand which parts of the image the model focuses on during classification can provide valuable insights.

3

Data Expansion

Collecting more diverse and challenging X-ray images can further enhance the model's robustness and generalization ability.

Conclusion

Through data preprocessing, model architecture, training, and fine-tuning, I achieved a significant improvement in performance. Future research will focus on further enhancing the model's accuracy and interpretability.