Probable Title:-

**Reconstructing 3D Shape from 2D Biplanar X-ray with Swin Transformers**

**Bir Fateh Singh, Ashmit Sheoran, Ansh Agrawal**

**Abstract**

This study explores the reconstruction of 3D shapes from 2D biplanar X-ray images using Swin Transformers. The referenced paper introduces a novel deep learning-based approach leveraging transformer architectures for medical imaging applications. Our work extends this research by refining the model pipeline, optimizing hyperparameters, and integrating additional pre-processing techniques to enhance reconstruction accuracy. By systematically analysing the performance of Swin Transformers in this context, we aim to provide further insights into their applicability and potential improvements.

**Introduction**

The reconstruction of 3D anatomical structures from 2D medical images is a crucial task in computer vision and medical imaging. Accurate 3D shape estimation is essential for applications such as surgical planning and diagnostic analysis. Traditional methods rely on statistical shape models and optimization techniques, but deep learning-based approaches, particularly transformers, have shown significant promise in recent years.

The referenced study employs Swin Transformers to reconstruct 3D shapes from biplanar X-ray images. The dataset used consists of medical images annotated with corresponding 3D models, enabling supervised training of the deep learning model. Our study builds upon this foundation, applying further refinements to the data pre-processing pipeline, model architecture, and evaluation metrics to enhance performance and generalizability.

**Methodology**

The original paper implements Swin Transformers to extract hierarchical features from 2D X-ray images, progressively reconstructing the 3D shape through a learned attention-based mapping. The approach involves:

* Preprocessing: Normalization, augmentation, and contrast enhancement of X-ray images.
* Feature Extraction: Utilizing Swin Transformer blocks to capture spatial dependencies and anatomical structures.
* Reconstruction: Predicting a 3D shape representation using learned feature embeddings and regression-based losses.

Our extension refines this methodology by:

* Experimenting with additional preprocessing methods to improve feature clarity.
* Fine-tuning transformer hyperparameters to optimize accuracy.
* Introducing an auxiliary loss function to improve structural consistency in predictions.
* Evaluating performance against alternative architectures, such as CNN-based methods, to establish comparative effectiveness.

**Results and Future Work**

The referenced study demonstrates that Swin Transformers achieve state-of-the-art performance in 3D reconstruction from 2D X-rays, surpassing traditional CNN-based models in structural accuracy. However, some challenges remain, including computational complexity and sensitivity to input noise.

Our preliminary results indicate that optimizing preprocessing techniques and hyperparameter tuning can further enhance the model’s accuracy and robustness. Future work will focus on:

* Reducing inference time while maintaining accuracy.
* Investigating multi-view fusion techniques to improve shape consistency.
* Exploring unsupervised or semi-supervised learning approaches to reduce reliance on annotated data.

**References**

1. Liu, Kuan, et al. *"Swin-X2S: Reconstructing 3D Shape from 2D Biplanar X-ray with Swin Transformers."* arXiv, 2025, <https://doi.org/10.48550/arXiv.2501.05961>.