

THE GENERATIVE GAP: GENERAL-PURPOSE VISION MODELS VS SPECIALISED NETWORKS FOR PTOSIS

Jaskaran Singh Bhangu, Chris Stewart, Rhadika Rewal, Parinita Swarnkar, Safa Elhassan, Gwyn Williams
Ophthalmology department, Singleton Hospital, Swansea Bay University Health Board, Swansea, Wales, UK.

Introduction

Can general-purpose AI (ChatGPT-5, Claude 3) replace domain-specific networks (UNet) for periorbital biometrics? Evaluated using the **Open-Source Periorbital Segmentation Dataset** (2,842 images).

Performance Data

Dice Similarity Coefficient (DSC) comparison:



Clinical Accuracy (MRDI)

Mean Absolute Error (MAE) measured in mm:

0.2mm

UNE
T

0.9mm

QWEN-
VL

Conclusion

Clinical accuracy requires error **< 0.5mm**. General-purpose models fail this threshold, requiring continued reliance on UNet for precise surgical planning and ptosis grading.

