



NARASARAOPETA ENGINEERING COLLEGE (AUTONOMOUS)
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
2025-2026

Batch Number	BG6
Team Members	A. Anitha(22471A0572) M. Kathyayani(22471A05A9) S. Harshini(22471A05C6)
Guide	M. Suresh M.Tech
Title	DeMT: Deformable Mixer Transformer for Multi-Task Learning of Dense Prediction
Domain/Technology	Deep Learning
Base Paper Link	https://arxiv.org/pdf/2301.03461
Dataset Link	https://cs.stanford.edu/~roozbeh/pascal-context/ https://www.kaggle.com/datasets/fatemehboloori/pascal-context-voc-2010 https://www.kaggle.com/datasets/soumikrakshit/nyu-depth-v2
Software Requirements	Browser: Any latest browser like Chrome Operating System: Windows 7 Server or later Python (COLAB)
Hardware Requirements	SystemType: Intel Core i5 or above RAM: 8 GB Number of cores:5 Number of Threads: 4
Abstract	This project introduces DeMT (Deformable Mixer Transformer), a powerful deep learning model designed for multi-task learning in computer vision. Traditional models use either CNNs or Transformers, each with strengths and weaknesses — CNNs capture fine local details, while Transformers handle global context. DeMT combines the best of both by using a deformable CNN encoder to focus on important image regions and a task-aware Transformer decoder to process multiple tasks like semantic segmentation, depth estimation, boundary detection, and more simultaneously. The model achieves state-of-the-art performance on popular datasets like NYUD-v2 and PASCAL-Context, outperforming existing approaches in accuracy and efficiency. This work shows that combining deformable convolution and query-based attention can lead to more accurate and task-aware multi-task models.

Signature of the student(s) Signature of the Guide Signature of the project coordinator