

04 Task organization

Figure 1 coordinate 4 | Subitems 13, 14, 15

Machine learning

13 Multitask

Several tasks -> shared model
Share learning across objectives
Keep separate task evaluations
Test each task versus specialists

14 Multimodal

Different modalities -> modeling
Link language and visual content
Conditioning is not joint sampling
Test cross-modal agreement

15 Joint generation

Noise -> corresponding outputs
Learn how outputs depend
Describe one shared sample
Test pairs and individual outputs

Computer vision

— Palette generalist model

Task + input -> restored image
Colorize, inpaint, uncrop, restore
Share conditional diffusion
Test each task, not their average

— UniDiffuser image and text

Text -> image; image -> text
Also supports paired sampling
Specify the sampling mode
Test alignment in each mode

— JointNet RGB and depth

Text + noise -> RGB-depth pair
Couple two diffusion branches
Target matching scene geometry
Test edges and pair consistency

Medical image analysis

— AortaDiff shared objectives

Plain CT -> synthetic CT / masks
Synthesis + segmentation tasks
Target shared anatomical identity
Test image / Dice / size error

— UniX reports and X-rays

Report -> X-ray; X-ray -> report
Generate across distinct formats
Target findings, not patient ID
Test omissions / contradictions

— MedSegFactory image / mask

Text + noise -> image-mask pair
Cross-condition two streams
Target matching anatomy / labels
Test boundaries / real-task utility