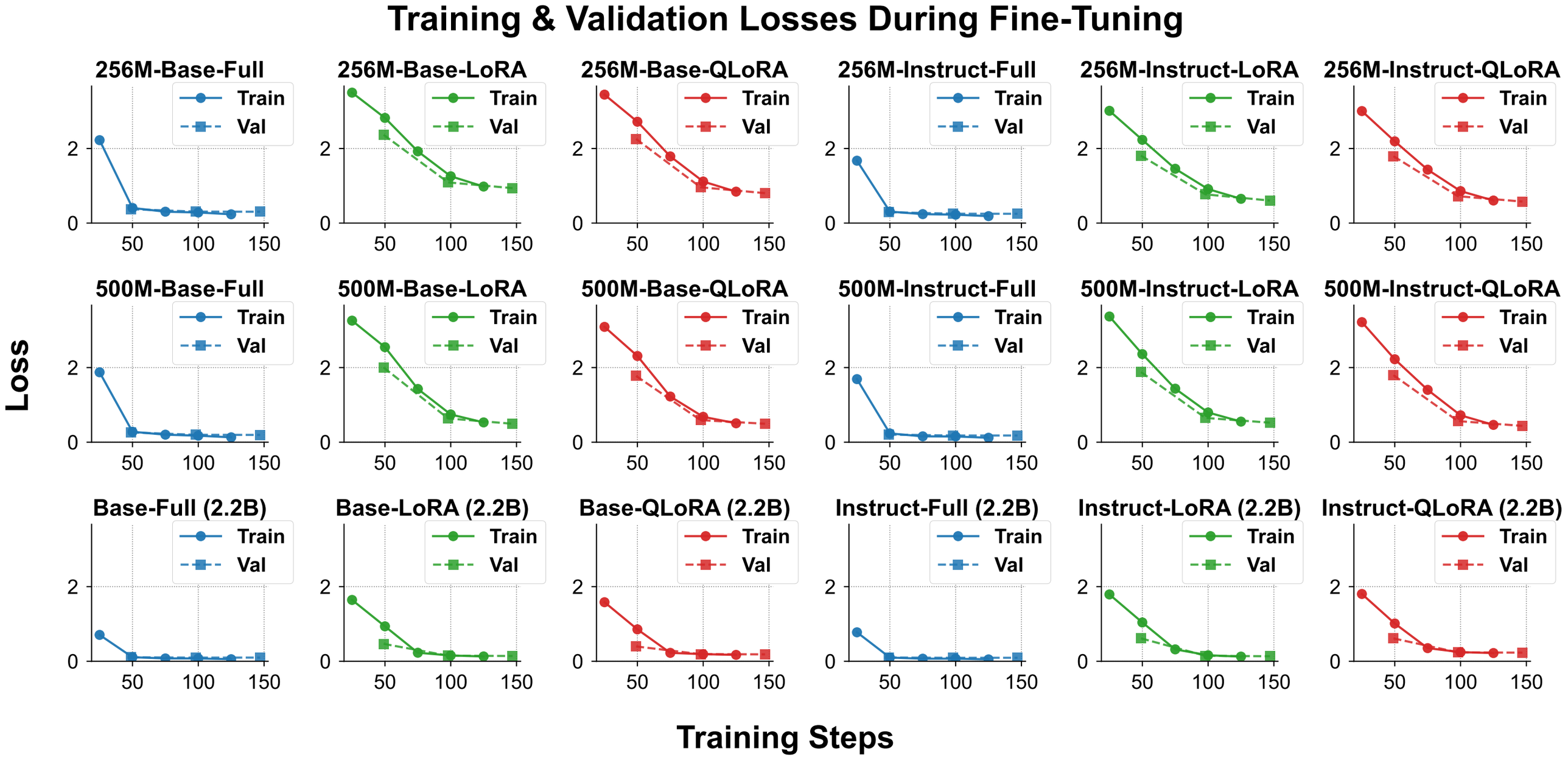
Training & Validation Losses During Fine-Tuning
256M-Base-Full
256M-Base-LoRA
256M-Base-QLoRA
256M-Instruct-Full
256M-Instruct-LoRA
256M-Instruct-QLoRA
500M-Base-Full
500M-Base-LoRA
500M-Base-QLoRA
500M-Instruct-Full
500M-Instruct-LoRA
500M-Instruct-QLoRA
Base-Full (2.2B)
Base-LoRA (2.2B)
Base-QLoRA (2.2B)
Instruct-Full (2.2B)
Instruct-LoRA (2.2B)
Instruct-QLoRA (2.2B)
Train
Val
Loss
Training Steps
0
2
50
100
150