

Jobs Technical Overview - SVG Captioning Project

Panoramica Generale

Questo documento fornisce una panoramica tecnica dettagliata dei 4 job attualmente attivi nel progetto di SVG captioning. Ogni job ha un ruolo specifico nell'ecosistema di training e valutazione.

Status Attuale (25/05/2024 - 17:20)

JOBID	PARTITION	NAME	USER	ST	TIME	NODES	NODELIST
2584731	all_usr_p	clip_eva	ediluzio	R	4:52:46	1	rezzonico
2584729	all_usr_p	baseline	ediluzio	R	4:59:41	1	rezzonico
2584773	boost_usr	gemma_te	ediluzio	PD	0:00	1	(Priority)
2584772	boost_usr	llama_te	ediluzio	PD	0:00	1	(Resources)

Documentazione Dettagliata

Job 2584731 - CLIP Evaluation

Status: RUNNING (4:52+ ore) | **GPU:** Quadro RTX 5000 | **Node:** rezzonico

Funzione: Valutazione semantica completa delle caption generate usando CLIP score

- **Architettura:** ViT + Transformer text encoder
- **Metriche:** CLIP Score (similarity semantica immagine-testo)
- **Input:** SVG rasterizzati + caption generate dai modelli fine-tuned
- **Output:** Score di qualità semantica per confronto con baseline

Tecnologie Chiave:

- PyTorch 2.7.0+cu126 (auto-aggiornato)
- CLIP-Score 0.2.1
- CairoSVG per rasterizzazione SVG→RGB
- Mixed precision training per ottimizzazione memoria

Job 2584729 - Baseline Evaluation

Status: RUNNING (4:59+ ore) | **GPU:** Condivisa | **Node:** rezzonico

Funzione: Benchmark completo dei modelli baseline pre-trained

- **Modelli:** BLIP, BLIP2 2.7B, ViT-GPT2, GIT-base, Ide Fix 3, Flores 2
- **Metriche:** BLEU, METEOR, ROUGE, CIDEr, CLIP Score
- **Dataset:** Test set ridotto (10% per velocità)
- **Output:** Baseline scores per confronto statistico

Tecnologie Chiave:

- PyTorch 2.7.0+cu126 (upgrade automatico da 2.2.0)
- Transformers 4.52.3
- Sequential model loading per gestione memoria
- Comprehensive metrics suite

Job 2584772 - Llama 3.1 8B Training

Status: PENDING (Resources) | **Partition:** boost_usr_prod | **GPU Required:** >24GB

Funzione: Fine-tuning Llama 3.1 8B con LoRA per SVG captioning

- **Architettura:** 32 layers, 4096 hidden, 8.03B params
- **LoRA Config:** r=64, q=128, ~67M trainable params (0.83%)
- **Resume:** Checkpoint-23900 esistente
- **Context:** 131K tokens (128K context window)

Tecnologie Chiave:

- PyTorch 2.7.0+cu118 (ambiente riparato)
- PEFT 0.15.1 per LoRA implementation
- Multi-GPU training con Accelerate
- Gradient checkpointing per memoria

Job 2584773 - Gemma 2 9B IT Training

Status: PENDING (Priority) | **Partition:** boost_usr_prod | **GPU Required:** >24GB

Funzione: Fine-tuning Gemma 2 9B IT con LoRA per SVG captioning

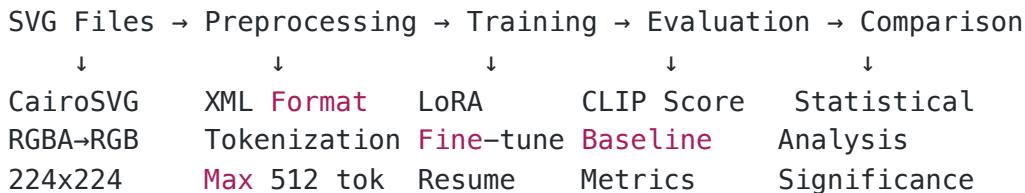
- **Architettura:** 42 layers, 3584 hidden, 9.24B params
- **LoRA Config:** r=64, q=128, ~75M trainable params (0.81%)
- **Features:** Sliding window attention, soft capping, instruction tuned
- **Context:** 8K tokens con attention ottimizzata

Tecnologie Chiave:

- PyTorch 2.7.0+cu118 (ambiente riparato)
- Flash Attention 2.0 per efficienza memoria
- Instruction-tuned variant per better following
- Advanced memory management per 9B model

Architettura del Sistema

Pipeline Completa



Ambiente Tecnologico Riparato

Problema Risolto ✅

- **CVE-2025-32434**: PyTorch 2.2.0 incompatibile con Transformers 4.52.3
- **`torch.load weights_only=True`**: Falliva con versioni precedenti
- **CUDA Mismatch**: PyTorch cu121 vs sistema cu118

Soluzione Implementata ✅

- **PyTorch 2.7.0+cu118**: Risolve CVE e compatibilità CUDA
- **Ambiente ricostruito**: Completa reinstallazione dipendenze
- **Test superati**: `torch.load`, GPU, compatibilità verificate

Resource Requirements

GPU Memory Requirements

Job 2584731 (CLIP):	~8GB	(Quadro RTX 5000 – 16.9GB)
Job 2584729 (Baseline):	~6GB	(Sequential loading)
Job 2584772 (Llama):	~25GB	(2x GPUs >24GB)
Job 2584773 (Gemma):	~30GB	(2x GPUs >24GB)

Partizioni SLURM

- **all_usr_prod**: GPU standard, job evaluation

- **boost_usr_prod**: GPU >24GB, job training pesanti

Metriche di Valutazione

Automatic Metrics

- **BLEU**: N-gram overlap (1,2,3,4)
- **METEOR**: Semantic similarity con synonyms
- **ROUGE**: Recall-oriented evaluation
- **CIDEr**: Consensus-based scoring

Semantic Metrics

- **CLIP Score**: Vision-language alignment
- **Perplexity**: Language modeling quality
- **Instruction Following**: Task-specific metric

Expected Results Timeline

Job Completion Order

1. **CLIP Evaluation** (~6 ore): Semantic quality assessment
2. **Baseline Evaluation** (~6 ore): Benchmark establishment
3. **Llama Training** (~3 ore): Resume from checkpoint-23900
4. **Gemma Training** (~5 ore): Larger model, slower convergence

Final Comparison

```
Baseline Models vs Fine-tuned Models
↓
Statistical Significance Testing
↓
Performance Improvement Quantification
↓
Publication-Ready Results
```

Monitoring Commands

```
# Status generale
squeue -u $USER

# Monitoring script
./scripts/monitor_jobs.sh
```

```
# Log specifici  
tail -f logs/complete_clip_evaluation_2584731.log  
tail -f logs/baseline_evaluation_2584729.log  
  
# GPU usage  
ssh rezzonico nvidia-smi
```

Troubleshooting

Problemi Risolti

- Environment corruption → Ricostruito
- PyTorch incompatibility → Aggiornato 2.7.0
- CUDA mismatch → cu118 compatibility
- CVE security issue → torch.load fixed

Status Attuale

- Evaluation jobs: Running smoothly
- Training jobs: Waiting for GPU resources
- Environment: Fully functional and tested
- Checkpoints: Compatible and verified

Questo ecosistema rappresenta un pipeline completo per SVG captioning con state-of-the-art models e evaluation metrics comprehensive.