GARUDEX LABS

Detailed Progress Report - November 2024

Ryan Madhuwala	Founder (ryan@garudexlabs.com)
Ronit Raj	Co-Founder (ronit@garudexlabs.com)
Samyak Choudhary	Core Member (samyak@garudexlabs.com)
Parv Mittal	Core Member (member@garudexlabs.com)

1 Executive Summary

November 2024 marked the inception of Garudex Labs, establishing our foundation in AI-driven surveillance solutions. Our team initiated groundbreaking research in multimodal learning models while setting up essential infrastructure and development frameworks.

2 Research and Development

2.1 Garudavyuha Model Architecture

- **Model Foundation**
  - Developed custom novel transformer-based architecture for image-text understanding
  - Implemented specialized attention mechanisms for surveillance contexts
  - Created architecture for combining image and text tokens efficient with SOTA benchmarks on it.
  - Designed feature extraction pipelines for real-time processing
- **Technical Specifications**
  - Base model size: 1.5B parameters
  - Context window: 4096 tokens
  - Processing speed: Initial benchmark at 20fps and 10-15 tokens/sec

2.2 Research Analysis

- **Literature Review**
  - Analyzed 25+ research papers on surveillance systems
  - Studied 10+ MLLM architectures for adaptation potential
  - Reviewed 8 anomaly detection frameworks
  - Examined 5 state-of-the-art image understanding systems
- **Algorithmic Development**
  - Created custom loss functions for surveillance tasks
  - Developed specialized attention mechanisms

  - Implemented efficient feature extraction methods
  - Designed novel data augmentation techniques

3 Technical Implementation

3.1 Infrastructure Setup

- **Development Environment**
  - Configured 2x NVIDIA T4 GPUs for training on kaggle
  - Configured local 2x NVIDIA RTX 3050 4GB/6GB for model running
  - Implemented distributed training pipeline
  - Set up version control with GitLab
  - Established continuous integration system
- **Data Pipeline**
  - Created efficient data preprocessing modules
  - Developed data validation frameworks
  - Set up automated data cleaning protocols

3.2 Model Training Progress

- **Initial Results**
  - Achieved 80.1% accuracy on VQAv2
  - Implemented batch processing capabilities
  - Optimized memory usage patterns
- **Optimization Efforts**
  - Applied gradient accumulation techniques
  - Implemented mixed precision training
  - Developed custom checkpointing system
  - Created automated performance monitoring

4 Future Roadmap

- **December 2024 Goals**
  - Complete video processing integration
  - Enhance model accuracy to 60%+ for video datasets
  - Implement long sequence understanding
  - Prepare for SIH 2024 presentation

- **Research Directions**
  - Explore advanced attention mechanisms
  - Investigate temporal analysis methods
  - Research efficient memory systems
  - Develop traffic management foundations