Forest-Chat: Adopting Vision-Language Models for Interactive Forest Change Analysis

James Brock, Ce Zhang, Nantheera Anantrasirichai
University of Bristol
ML4EO 2026

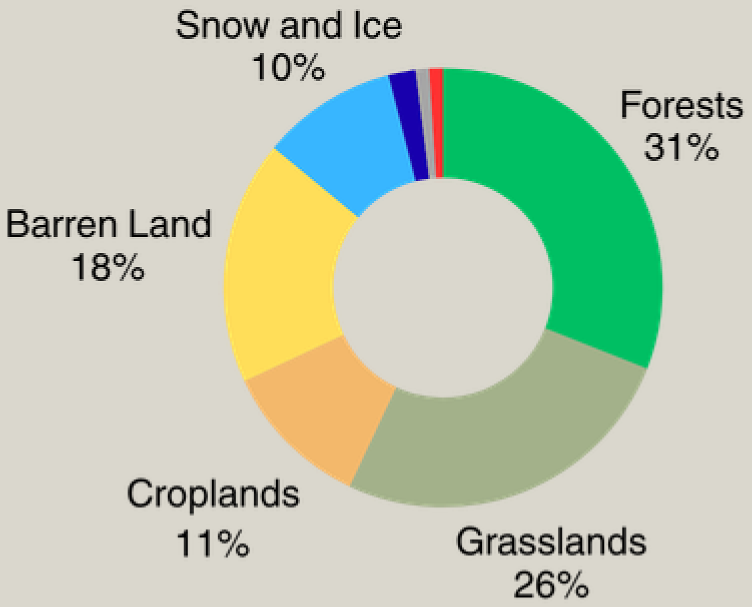
UK Research and Innovation

PrO·AI

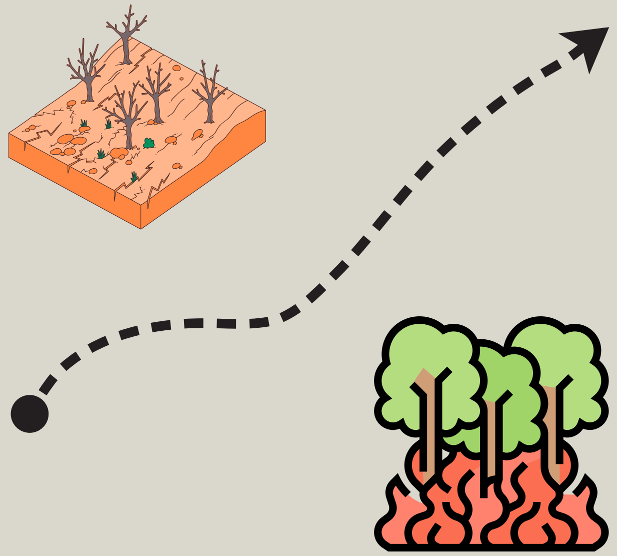
University of BRISTOL

Forest monitoring increasingly relies on AI and remote sensing

Forests dominate land cover and are critical for climate regulation



Increasing changes: deforestation, agriculture, wildfires, regeneration



Remote sensing enables large-scale monitoring

Challenge: interpreting what changed and why is still difficult at scale

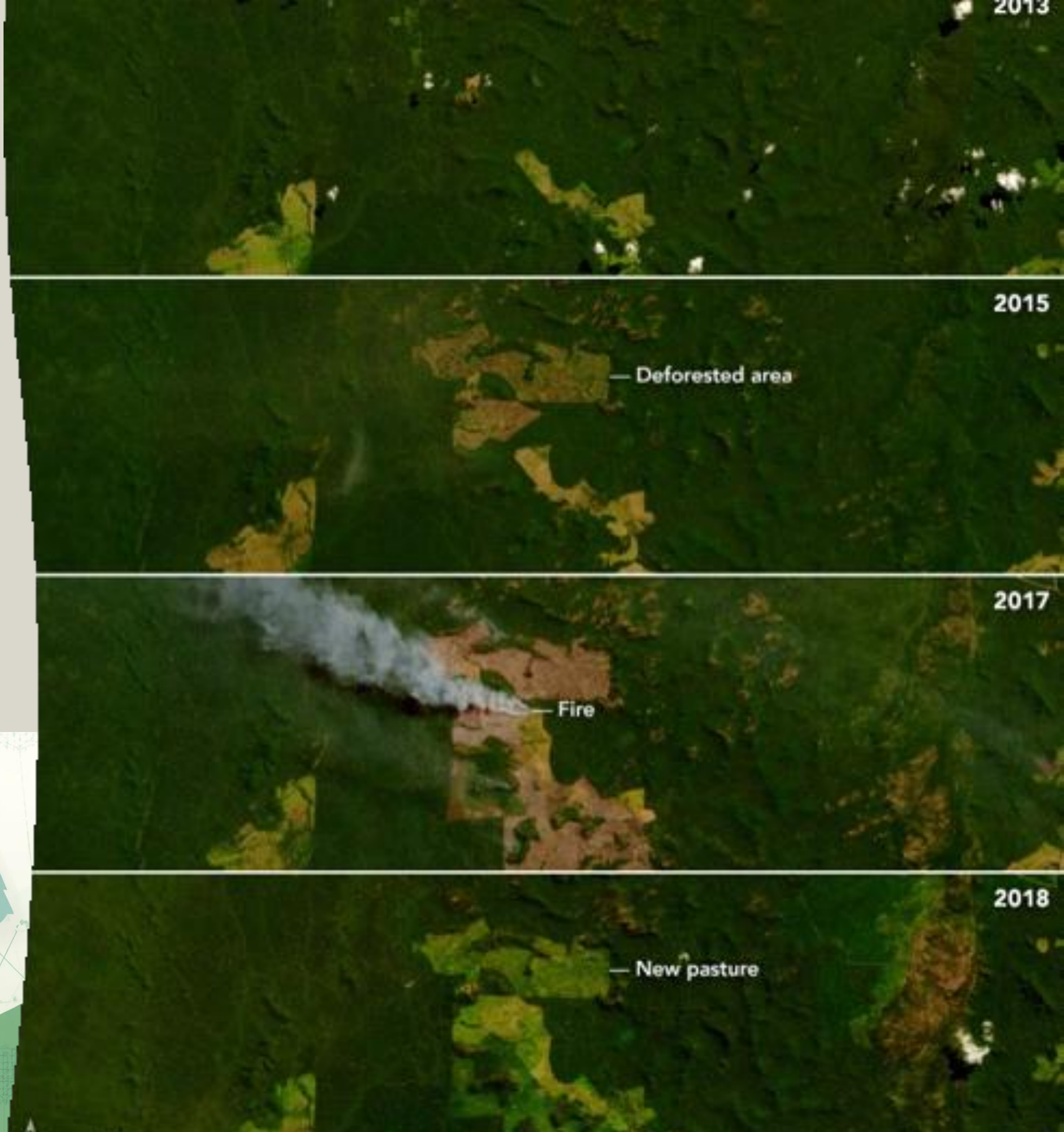
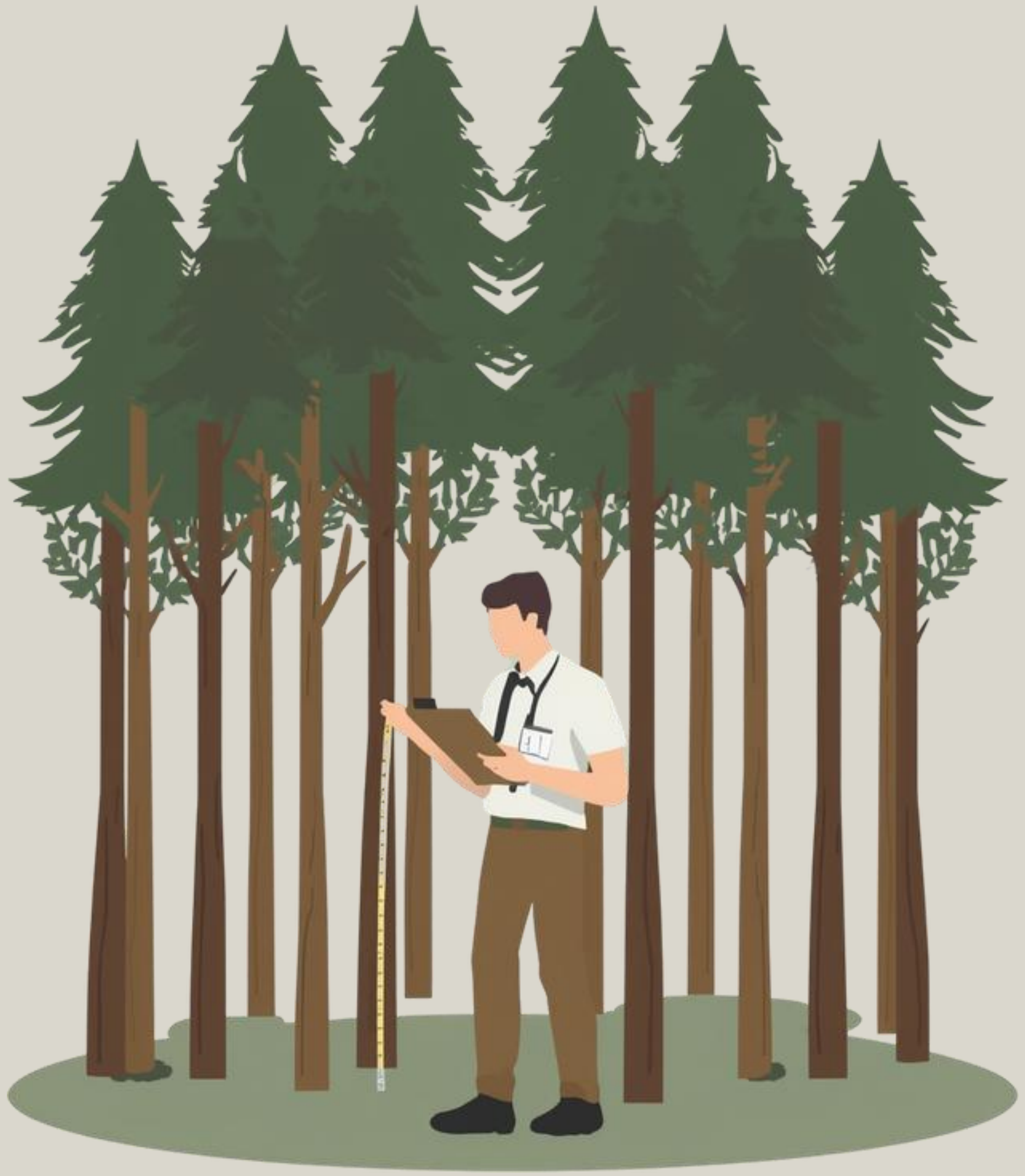


Image Credit: [1]

Existing systems don't effectively serve foresting monitoring workflows

Current Limitations

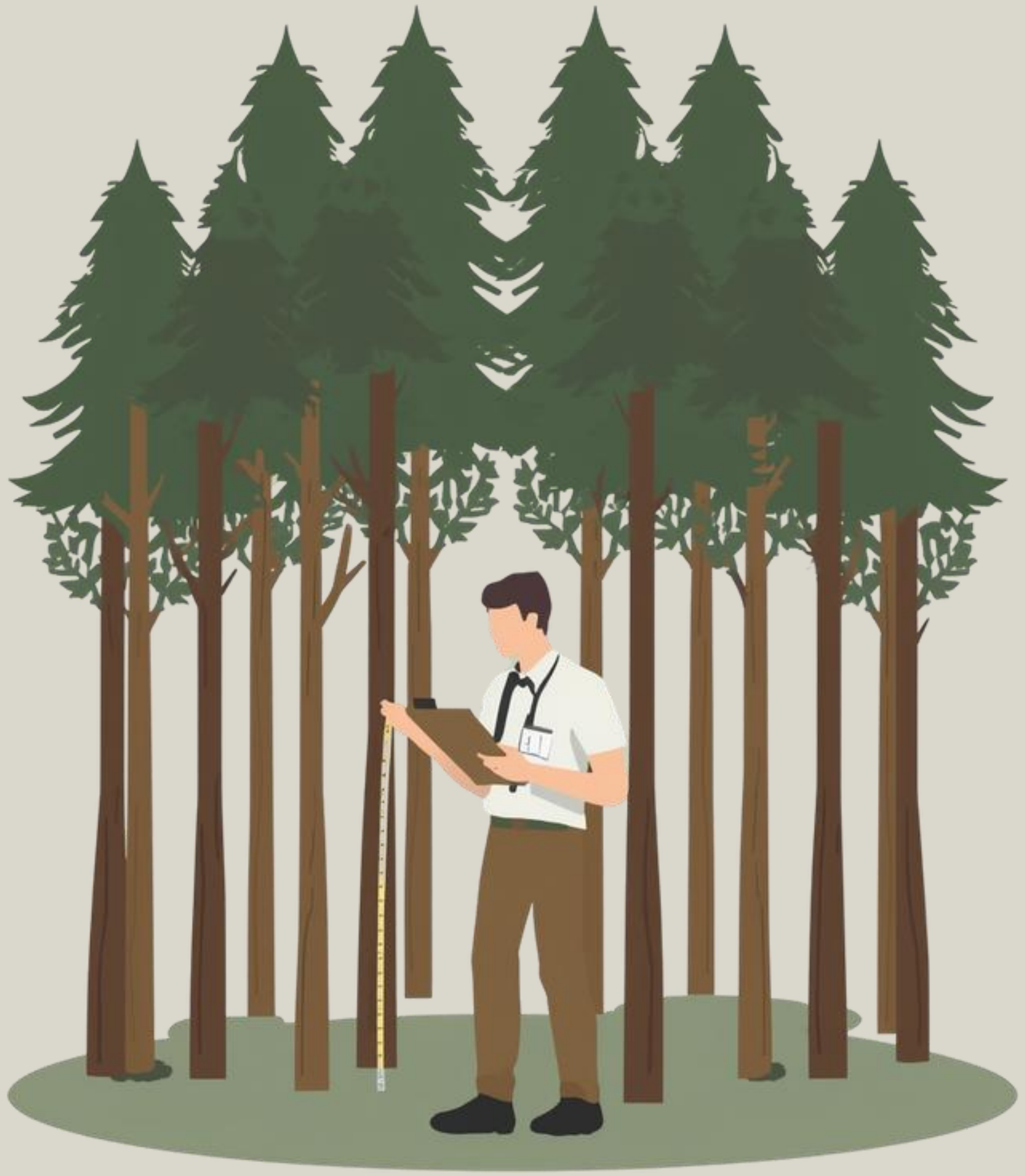
- Fragmented & task-specific
- Weak semantic understanding and lack of reasoning
- Data limitations
- Poor generalisation to domain shifts



Existing systems don't effectively serve foresting monitoring workflows

Current Limitations

- Fragmented & task-specific
- Weak semantic understanding and lack of reasoning
- Data limitations
- Poor generalisation to domain shifts



What's Needed Next

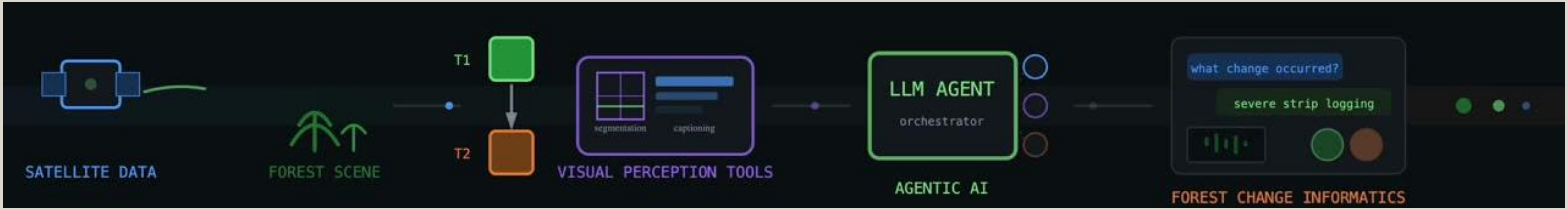
- Unified systems for detection + interpretation
- Richer multimodal understanding
- Increased dataset quantity + quality
- Robust, adaptable, interactive workflows with flexible orchestration

"Not just better models - but systems that reason, interact, and coordinate."

Idea: An agentic approach to interactive forest change analysis

Input: Bi-temporal satellite imagery of forest scenes

Reasoning & coordination: decompose user queries into executable tasks



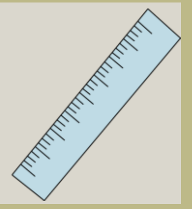
Perception: Vision-language understanding of change dynamics incorporating what, where, and why?

Output: Flexible, interpretable understanding of forest change dynamics

→ **Generating bespoke forest scene insights from satellite imagery**

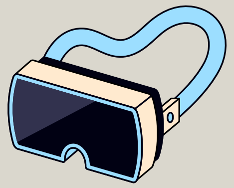
Research objective: Assess VLM-based agent capabilities for interactive forest change analysis

Benchmarking



Joint change detection and captioning against SOTA models on two datasets

Zero-shot capabilities



Explore vision foundation model performance in zero-shot settings

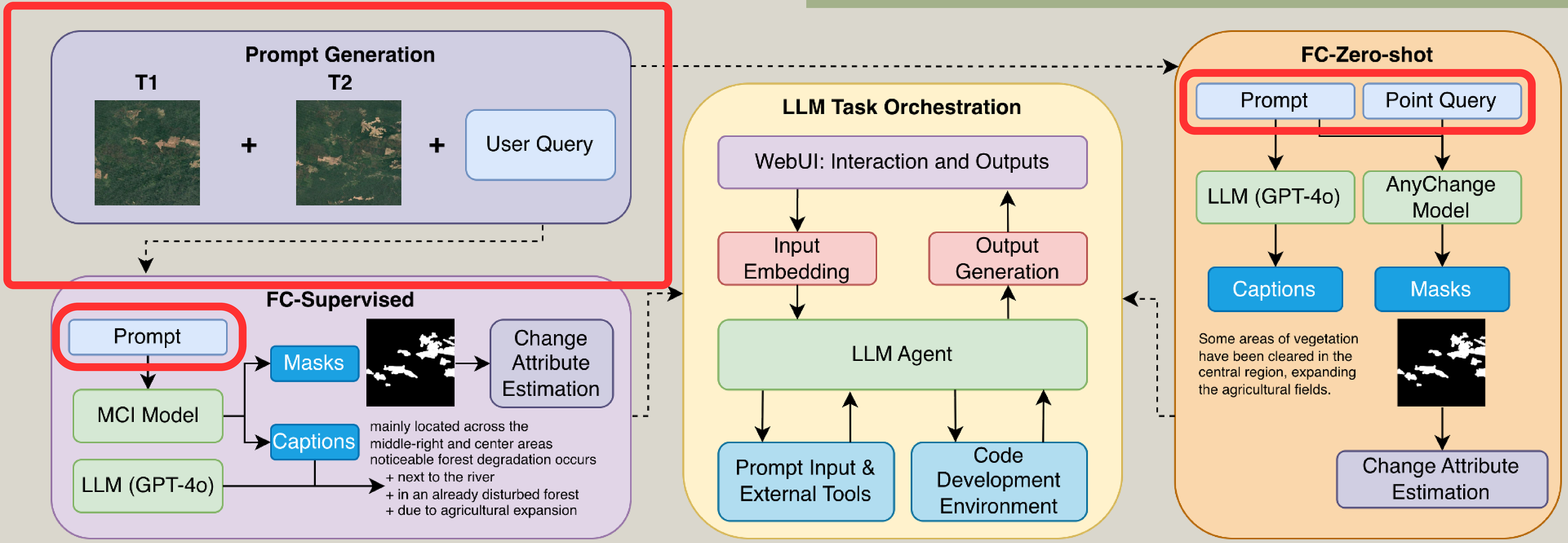
Cross-domain adaptation

Analyse how well learned features transfer across different data domains

Conversational utility

Validate vision-language agent robustness for forest change analysis

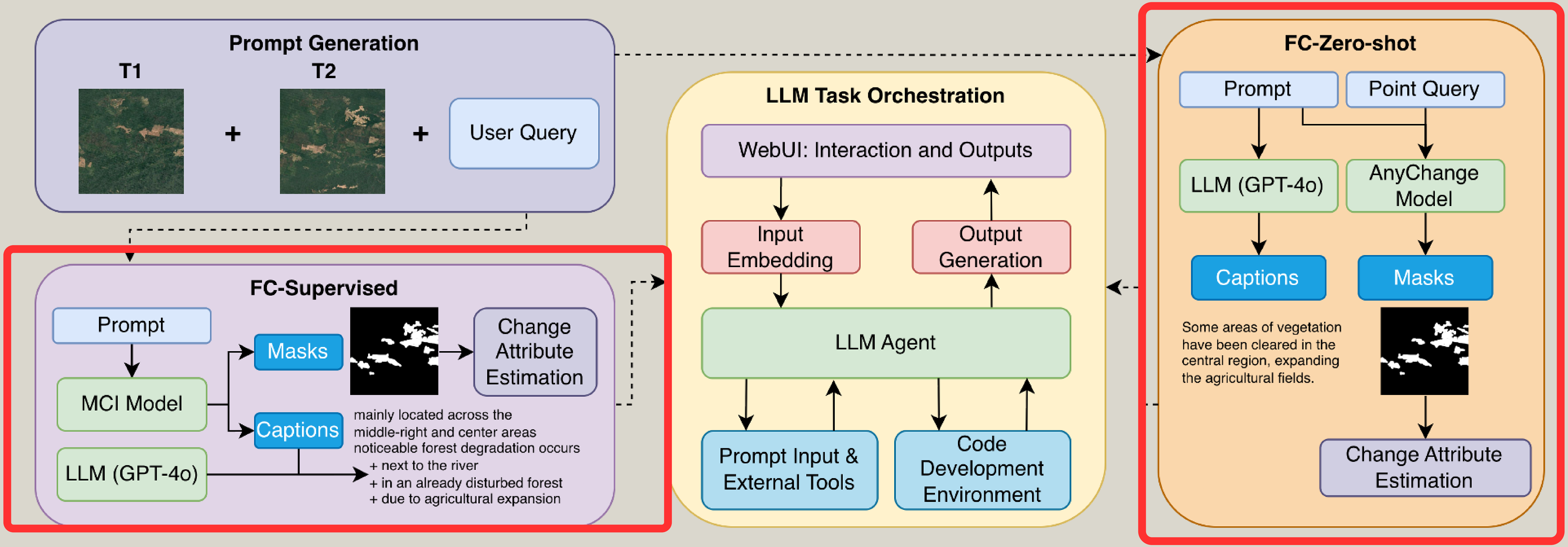
Forest-Chat agent overview



Input

- Bi-temporal satellite images
- Text-based user query
- Interactive UI

Forest-Chat agent overview



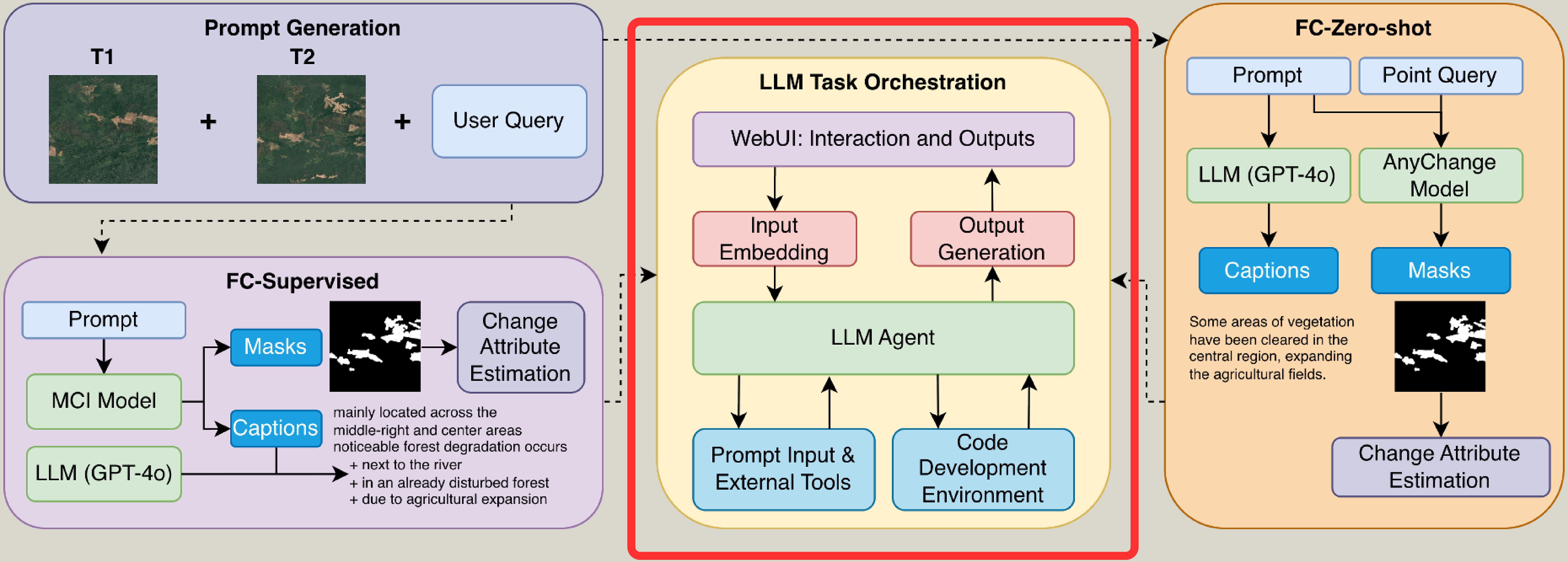
Input

- Bi-temporal satellite images
- Text-based user query
- Interactive UI

Perception

- Unified CDC model (MCI[3])
- Zero-shot CD (AnyChange[4])
- Zero-shot CC and caption refinement (GPT-4o[5])

Forest-Chat agent overview



Input

- Bi-temporal satellite images
- Text-based user query
- Interactive UI

Perception

- Unified CDC model (MCI[3])
- Zero-shot CD (AnyChange[4])
- Zero-shot CC and caption refinement (GPT-4o[5])

Orchestration

- LLM-agent task understanding
- Code orchestration framework
- External tool utilisation

Remote sensing image change interpretation lacks adequate benchmark datasets

Dataset	Main Task(s)	Key Characteristics	Labels Available	Resolution	Samples
Forest-Change[2, 7]	Change Detection + Captioning	Small, imbalanced, subtle deforestation, human + rule-based annotation	Masks + 1:4 Human to Rule-based Annotated Captions	30m	334
LEVIR-MCI-Trees[3]	Change Detection + Captioning	Urban object change masks, tree-focused captions, clearer structures	Masks + 5 Human Annotated Captions	0.5m	2305
JL1-CD-Trees[6]	Change Detection (domain transfer)	Tree cover gain + loss, domain shift, seasonal & atmospheric noise	Masks only	0.5 - 0.75m	408

- three noted areas of loss in the top left top right and centre on the edge of previous clearings new loss is moderate whilst overall loss is considerable
- slight forest degradation is noted mainly located in the top-left area
- concentrated in the top-left region small amounts of forest loss are apparent

Finding 1: Robust performance across multiple tasks in supervised and zero-shot settings

Finding 2: MLLM's can be utilised for post-captioning domain knowledge injection

Generic prompting

- Capture key changes
- Generic LLM style
- Can hallucinate

Style-guide prompting

- Reflects original style
- Few-shot examples reduce hallucinations

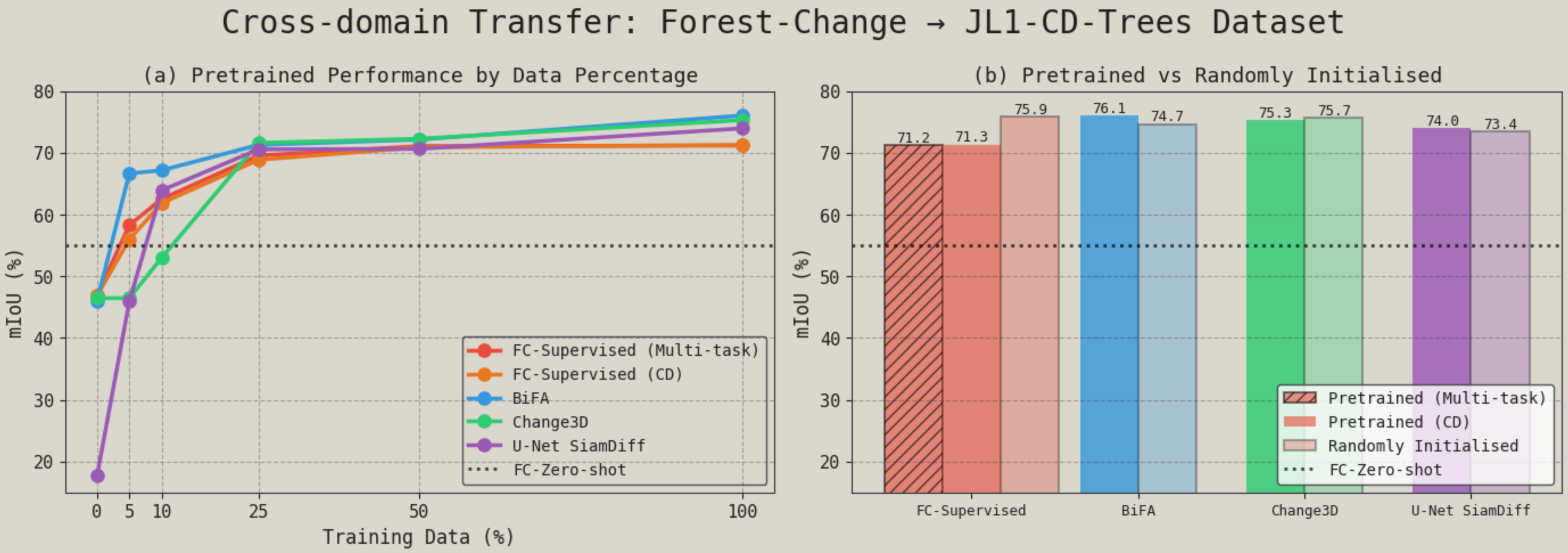
Refined captions

- Adds missing geographical info
- Can artificially reduce scores
- Corrects mistakes

Finding 3: Minimal data requirements for cross-domain adaptation

Zero-shot performance sensitive to object-specific ground truth masks

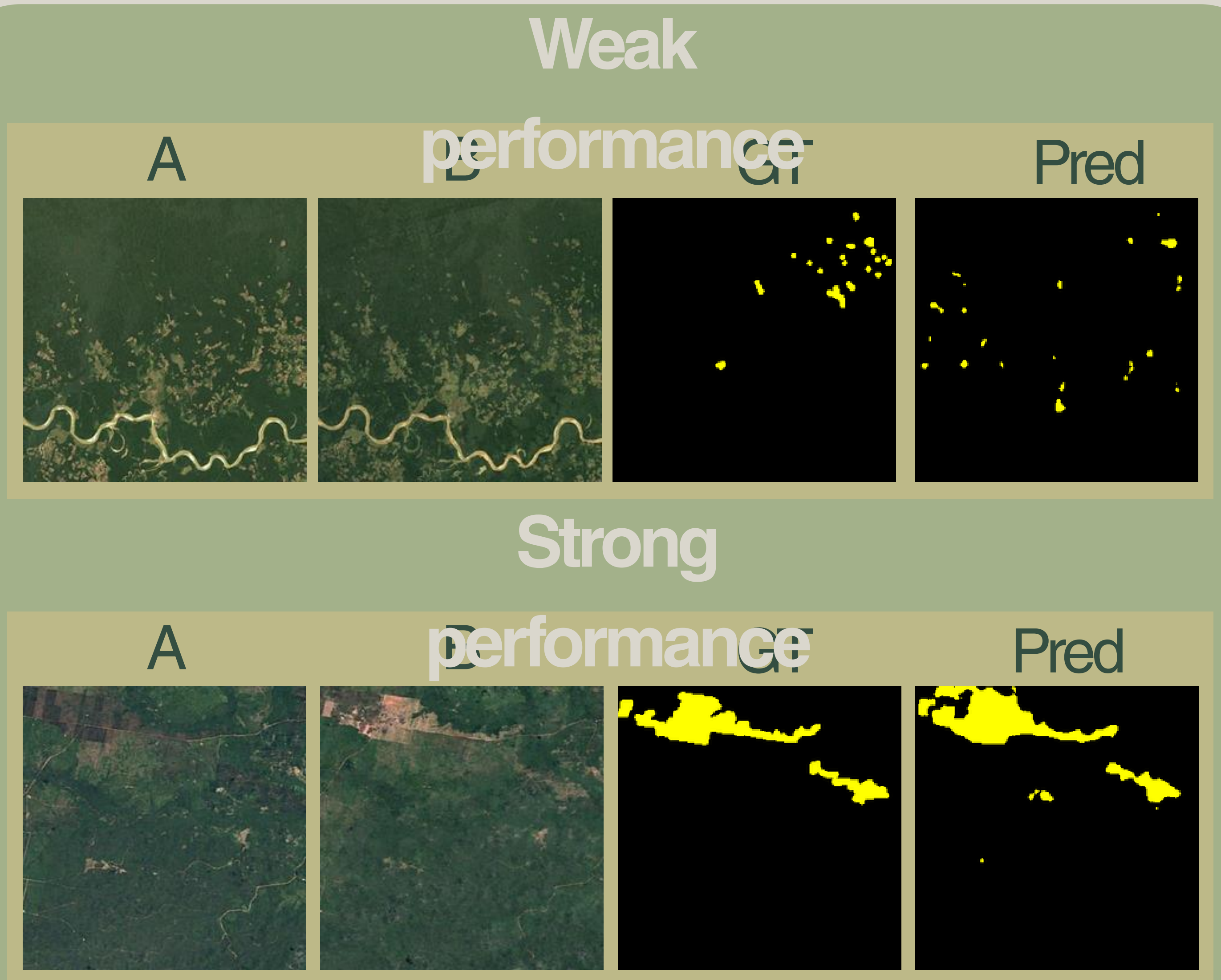
FC-Zero-shot Baselines	
Dataset	**MIoU**
Forest-Change	**60.15**
LEVIR-MCI-Trees	**47.32**
JL1-CD-Trees	**54.98**



Detection limits, data gaps, and system fragility constrain real-world impact

Remaining Challenges

- Struggles with small, subtle, and fragmented change patches
- Dataset size, quality, and task support limits model potential
- Scaling system to incorporate additional capabilities is non-trivial

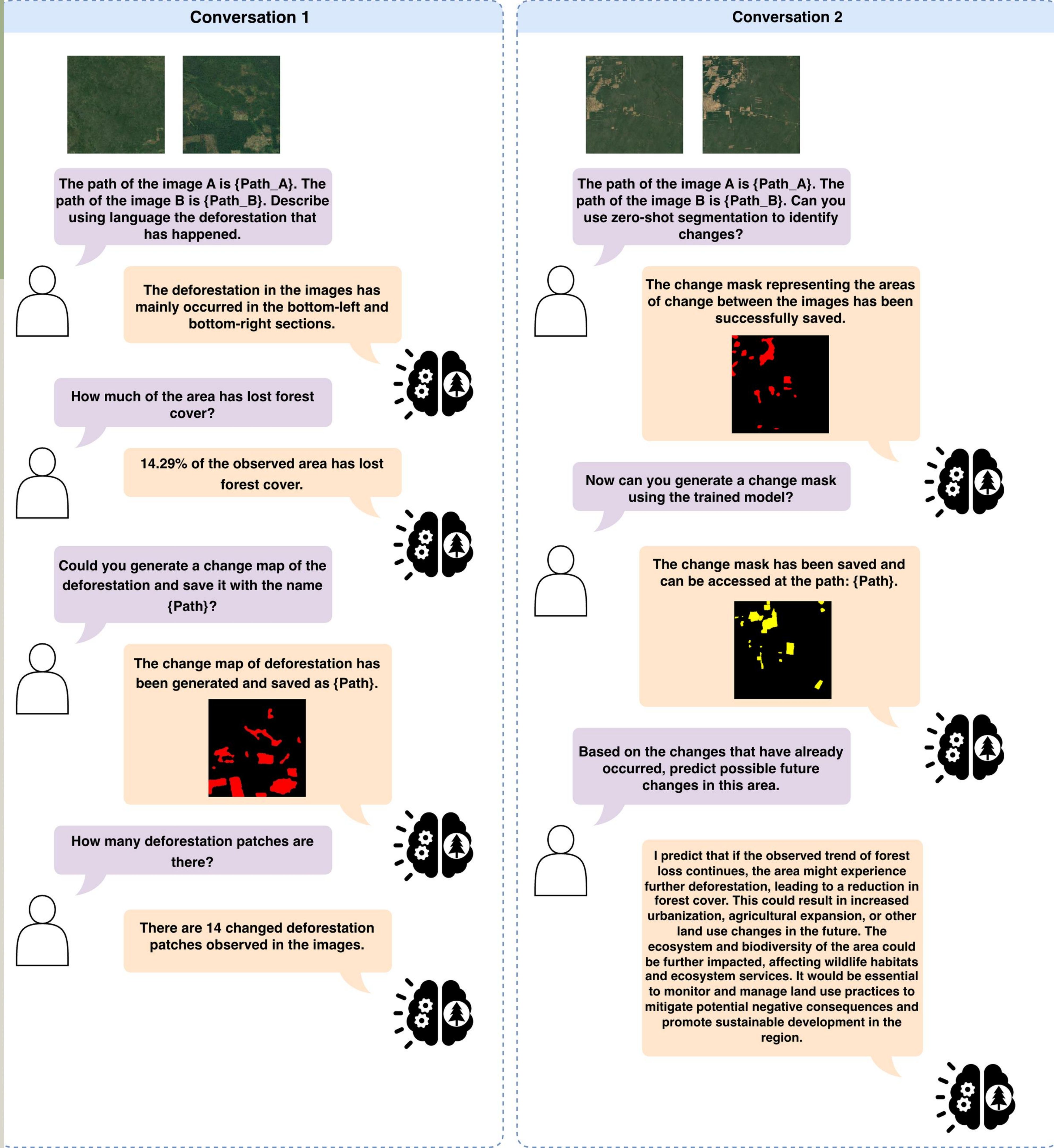


Promising Directions

- Auxiliary data and localisation architectures for small patches
- Larger, multi-modal datasets with richer annotations
- Task-aware retrieval and expanded toolsets

Forest-Chat: a practical integration study into interactive forest change analysis

- VLM and agentic integration yields a capable system for supervised and zero-shot forest change analysis
- Ecological RS exposes domain-specific challenges in detection, captioning, and agent reasoning
- Forest-Change and Forest-Chat provide a reproducible baseline for future work



References

1. NASA Earth Observatory. (2019). Making sense of Amazon deforestation patterns. Retrieved July 9, 2025, from https://earthobservatory.nasa.gov/images/145888/making-sense-of-amazon-deforestation-patterns?src=ve
2. Hewarathna, A. I., Hamlin, L., Charles, J., Vigneshwaran, P., George, R., Thuseethan, S., Wimalasooriya, C., & Shanmugam, B. (2024). Change Detection for Forest Ecosystems Using Remote Sensing Images with Siamese Attention U-Net. Technologies, 12(9), 160. https://doi.org/10.3390/technologies12090160
3. Liu, C., Chen, K., Zhang, H., Qi, Z., Zou, Z., & Shi, Z. (2024). Change-agent: Towards interactive comprehensive remote sensing change interpretation and analysis. IEEE Transactions on Geoscience and Remote Sensing.
4. Zheng, Z., Zhong, Y., Zhang, L., & Ermon, S. (2024). Segment any change. arXiv preprint arXiv:2402.01188.
5. Hurst, A., Lerer, A., Goucher, A.P., Perelman, A., Ramesh, A., Clark, A., Ostrow, A.J., Welihinda, A., Hayes, A., Radford, A. and Mądry, A., 2024. Gpt-4o system card. arXiv preprint arXiv:2410.21276.
6. Liu, Z., Zhu, R., Gao, L., Zhou, Y., Ma, J. and Gu, Y., 2025. JL1-CD: A new benchmark for remote sensing change detection and a robust multi-teacher knowledge distillation framework. arXiv preprint arXiv:2502.13407.
7. Brock, J., Zhang, C., & Anantrasirichai, N. (2026). Forest-Chat: Adapting vision-language agents for interactive forest change analysis. Ecological Informatics, 95, 103741. https://doi.org/10.1016/j.ecoinf.2026.103741

Acknowledgements

Authors wish to acknowledge and thank the financial support of the UK Research and Innovation (UKRI) [Grant ref EP/Y030796/1] and the University of Bristol.

Connect and explore further!

LinkedIn

Full Paper

GitHub Repo