

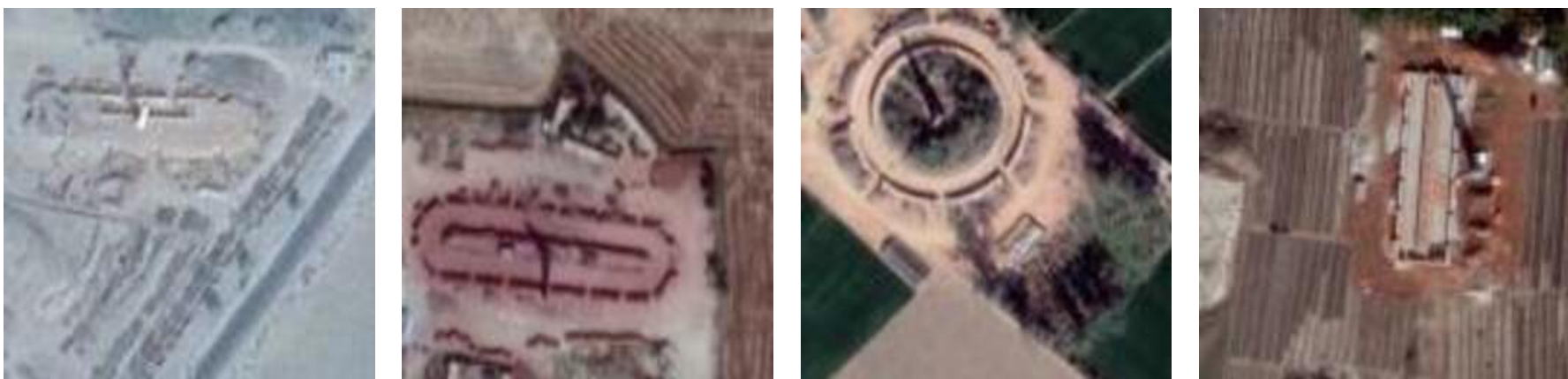
Foundation Models for Mapping Emission Sources and Acute Respiratory Infection (ARI) Hotspots

1. Motivation

- **Health burden:** Acute Respiratory Infections (ARI) are a leading cause of child morbidity in LMICs [1].
- **Emission sources:** Unregulated brick kilns and industrial chimneys release SO₂, NO₂, CO, and particulates—yet their *spatial link* to ARI is under-quantified at scale.
- **Data gap:** We lack high-resolution, region-wide maps connecting emission sources to health outcomes.
- **Opportunity:** Combine foundation models + satellite imagery with geo-coded DHS surveys to quantify risk and target interventions.

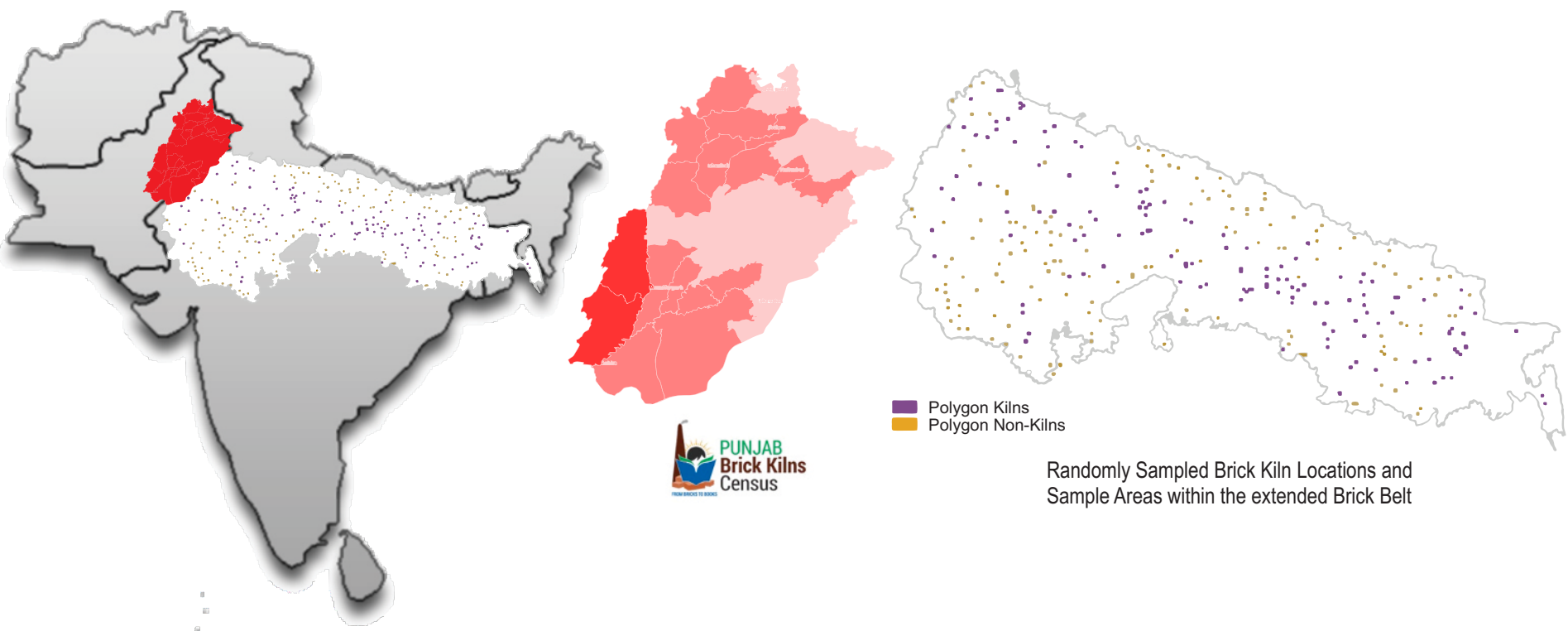
2. Challenges

- Bottleneck: Absence of spatially explicit data.
- Huge variations in satellite imagery.



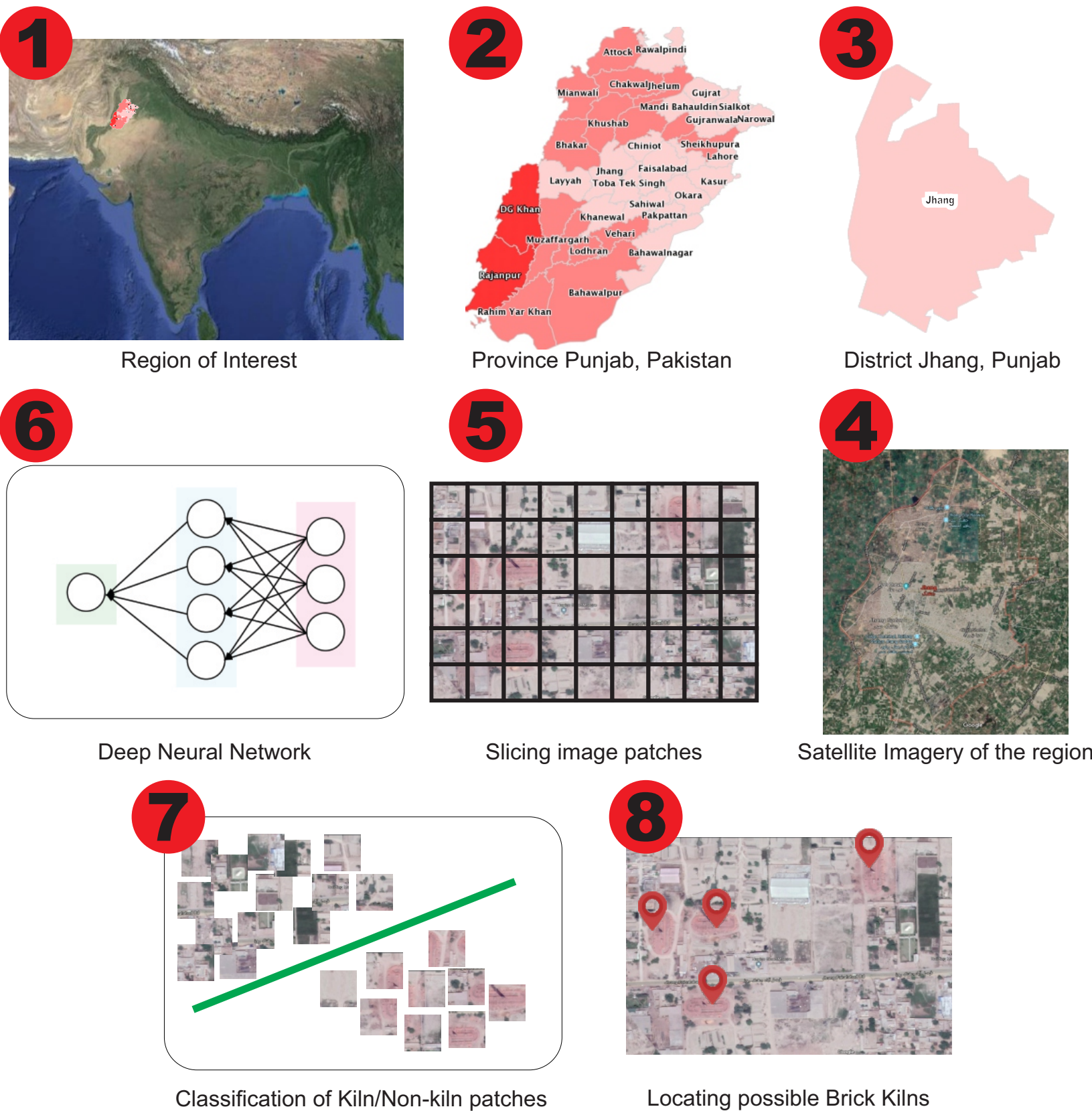
3. Existing Strategies & Limitations

- Field surveys and manual localization are both tedious and lack scalability [2].
- Data obtained manually from satellite imagery is very sparse (320 kilns only) [3].

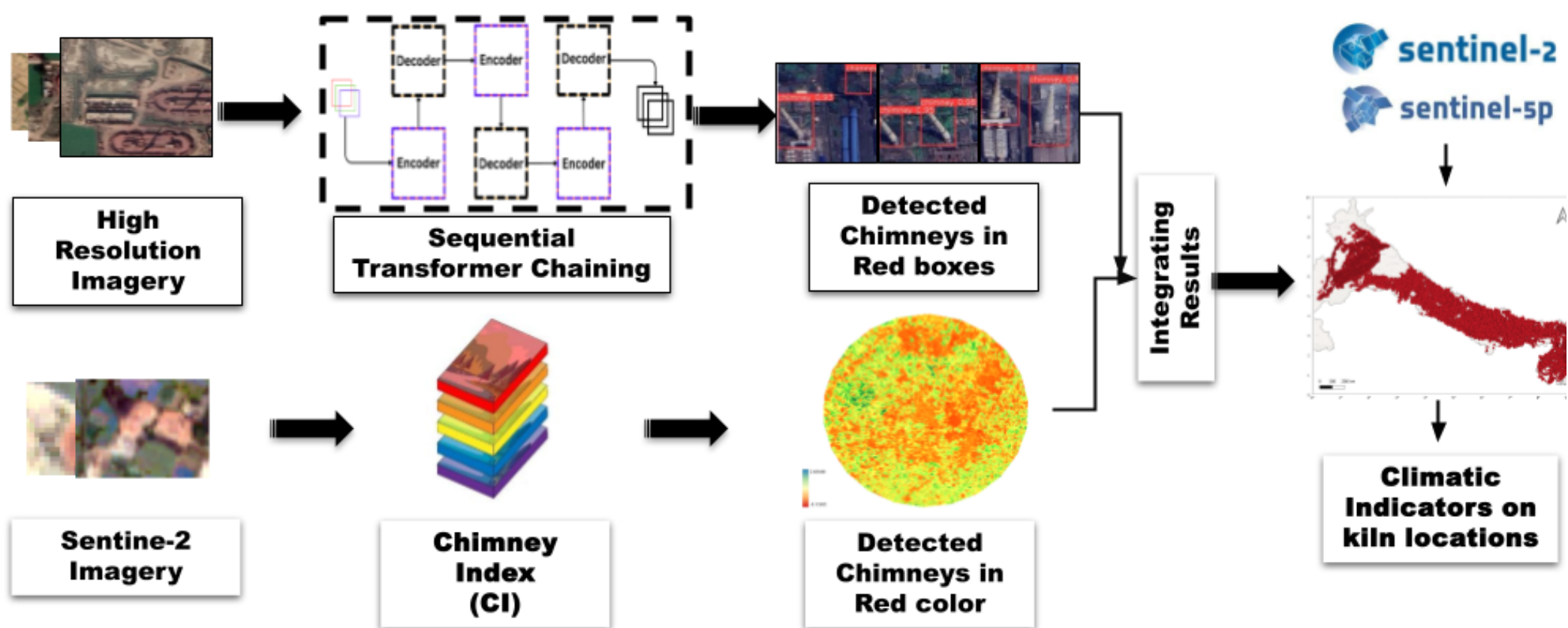


4. Emission Sources Detection Framework

- A classical machine-learning approach to localize brick kilns in remote-sensing imagery.



5. Proposed Architecture



6. Quantitative Results on Pilot Dataset

Table 1. Correlation between ARI cases and environmental indicators. Top-2 correlations are highlighted.

Variable	CO ₂ _mean	NO ₂ _mean	SO ₂ _mean	Mean Temp. (°C)
NO_ARI_CASES	0.005	0.031	0.170	0.001

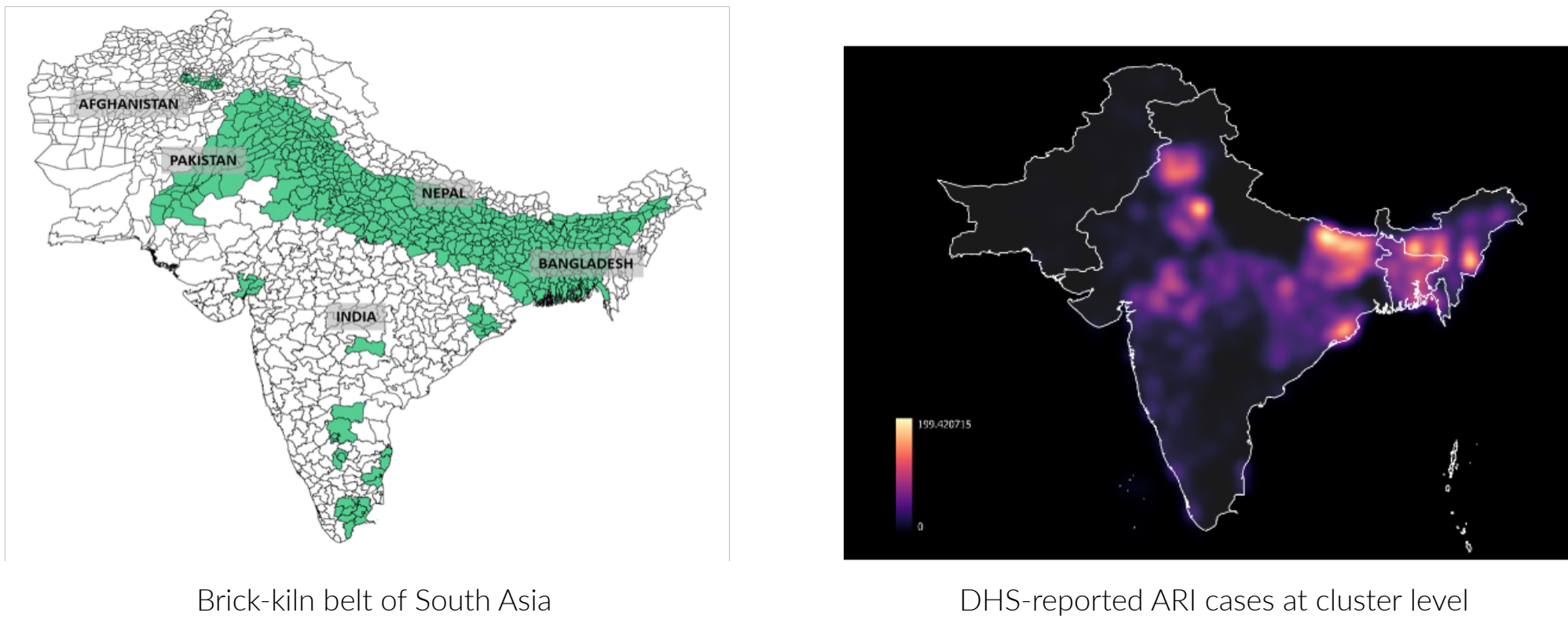
Table 2. Performance of chimney detection methods. Top-2 accuracies are highlighted.

Method / Stage	Accuracy (%)
Chimney Index (threshold = 0.4)	
Patparganj Industrial Area, Delhi (India)	79.17
Sundar Industrial Estate, Lahore (Pakistan)	68.42
Tongi Industrial Estate, Dhaka (Bangladesh)	70.00
Sequential Transformer Chaining	
Initial Vision Transformer Filtering	60.00
Secondary Remote CLIP Filtering	85.00
Tertiary Remote CLIP Refinement	95.00

Table 3. ARI cases prediction using multiple climatic and air-quality indicators.

Model	Loss Function	Validation MSE
BERT (baseline)	MSE	0.860
TabTransformer (Our implementation)	MSE	0.145

7. Qualitative Results



Takeaways & QR

- **Weak Associations:** Emission sources spatially correlate with pediatric ARI, with SO₂ most indicative.
- **Tooling works:** Foundation-model chaining yields a precise, scalable chimney inventory.
- **Actionable:** The pipeline supports targeted health and regulatory interventions in high-risk zones.



Paper, code, and data pointers

References

- [1] World Health Organization. *Environmental Risk Factors and the Burden of Disease*. 2018. URL: <https://www.who.int/about/accountability/results/2018-2019>. Accessed: 2025-08-05.
- [2] Punjab Brick Kiln Census, URL: http://202.166.167.115/brick_kiln_dashboard/
- [3] Boyd, Doreen S., et al. "Slavery from Space: Demonstrating the Role for Satellite Remote Sensing to Inform Evidence-Based Action related to UN SDG number 8." *ISPRS J. Photogrammetry and Remote Sensing* 142, 2018: 380–388.