

BiSign: Bidirectional Cross-Lingual Transfer Learning for Sign Language Recognition

Ankit Utkarsh Hota

UG Student, Department of Computing Technologies, SRM Institute of Science & Technology, Kattankulathur, Chennai, Tamil Nadu, India

Yash Ranjan

UG Student, Department of Computing Technologies, SRM Institute of Science & Technology, Kattankulathur, Chennai, Tamil Nadu, India

Brindha R

Assistant Professor, Department of Computing Technologies, School of Computing, SRM Institute of Science & Technology, Kattankulathur, Chennai, Tamil Nadu, India

Dr. Malarselvi G

Assistant Professor, Department of Computing Technologies, School of Computing, SRM Institute of Science & Technology, Kattankulathur, Chennai, Tamil Nadu, India

Dr. Ramesh S

Assistant Professor, Department of Computing Technologies, School of Computing, SRM Institute of Science & Technology, Kattankulathur, Chennai, Tamil Nadu, India

Abstract:

Sign language recognition (SLR) systems suffer from severe data scarcity outside of American Sign Language (ASL). We present BiSign, a cross-lingual transfer learning framework that trains a single shared encoder on ASL and Indian Sign Language (ISL) jointly, using a language-ID conditioning token and NT-Xent contrastive loss to align embedding spaces across languages. Evaluated on the WLASL (1,575 classes) and INCLUDE (239 classes) benchmarks over three random seeds, our key results are: (1) a 5-shot ASL-pretrained model achieves $53.9 \pm 0.8\%$ ISL Top-1 accuracy versus $0.6 \pm 0.4\%$ for 5-shot scratch training (+52.9 pp); (2) a 10-shot pretrained model reaches $66.7 \pm 0.2\%$ versus 57.1% fulldata scratch (+9.6 pp with 14× less labelled data); (3) zero-shot transfer achieves 32.0%, which is 77× above random chance. Ablations confirm the language-ID token contributes +9.3 pp and the shared encoder contributes +16.2 pp over an ISL-only separate encoder.

Keywords:

Sign language recognition, cross-lingual transfer learning, few-shot learning, transformer, MediaPipe, WLASL, INCLUDE.