

# Couples Behavior Modeling and Annotation Using Low-Resource LSTM Language Models

Shao-Yen Tseng<sup>1</sup>, Sandeep Nallan Chakravarthula<sup>1</sup>, Brian Baucom<sup>2</sup>, Panayiotis Georgiou<sup>1</sup>

<sup>1</sup>USC, Department of Electrical Engineering, SCUBA Lab

<sup>2</sup>The University of Utah, Department of Psychology

## Motivation & Introduction

- Studies in couples therapy use behavior codes
- Manual annotations of behavior codes are
  - Expensive
  - Time-consuming
  - Low inter-annotator agreement
- We investigate the application of RNNs to
  - Capture behavior trajectories
  - Estimate ratings from fusion of frame-level behavior

## Data

UCLA/UW Couple Therapy Research Project

- 134 real couples with marital issues interacting
- Couples discuss specific topic for 10 min each
- Behaviors rated by multiple annotators
- Manual transcription into text

### Challenges

- One global rating for entire session
- Learning with noisy labels
- Learning with limited data

## Methodology

### Proposed Architecture

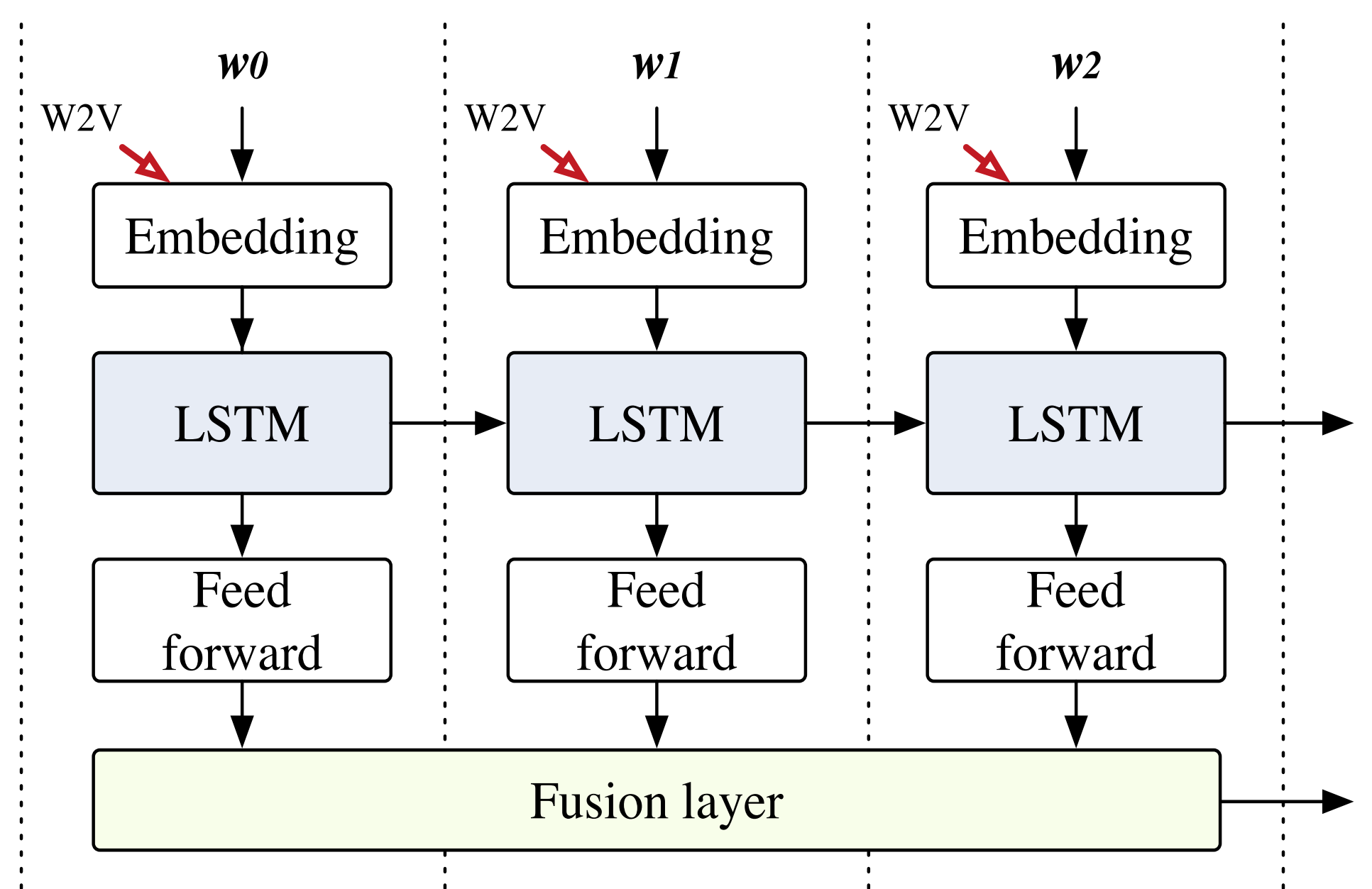
- Recursive neural network with 3 hidden layers
- Fusion layer for session label prediction

### Out-of-Domain Word Representations

- OOD knowledge through word embeddings
- Pre-trained on 4B words from Google News

### Fusion Layer

- RNN predicts behavioral ratings on word sequences
- Fuse local predictions to obtain global ratings
- RBF-Kernel SVR on median of frame-level scores



## Results & Discussion

Annotator Configuration	Krippendorff's Alpha	
	No w2v	w2v+joint
All human annotators	0.821	
Random with machine predictions	0.7611	0.7739
Outlier with machine prediction	0.7997	<b>0.8249</b>

### Conclusion

- Word vectors help train RNNs with limited data
- Trained models for obtaining continuous metrics of behavior
- Able to predict actual behavior ratings through top-layer fusion comparable to humans

### Future Work

- Transfer learning from different behavior codes
- Fusion from different modalities