

# Contexto

En esta presentación se comparan los papers proporcionados, que tratan de los mecanismos de atención. Las comparaciones se centrarán en las referencias 2-4.

# Metodología de búsqueda

Texto encontrado en Google Scholar.

# Términos Preliminares

- **SoftMax function:** Conserva el orden de los valores de entrada, de menor a mayor, y los traduce en números entre 0 y 1 que suman 1.
- **Self-attention:** Mecanismo de atención que relaciona distintas posiciones de una secuencia única para calcular una representación de la secuencia. Se utiliza para la comprensión lectora, el resumen abstracto, la implicación textual y el aprendizaje de representaciones de oraciones independientes de la tarea.
- **End-to-end memory networks:** NN con un modelo de atención recurrente sobre una memoria externa posiblemente grande
- **Transducción:** transformación de un tipo de señal o energía en otra de distinta naturaleza.
- **Computer Vision:** Campo de la inteligencia artificial que capacita a las computadoras para interpretar y comprender el mundo visual.
- **SER: Speech Emotion Recognition.** Campo de estudio que trata de inferir emociones humanas a partir de señales de habla.
- **RFE: Recursive Feature Elimination.** método de selección de características que ajusta un modelo y elimina la característica (o características) más débiles hasta que se alcanza el número especificado de características.
- **ERC:** Emotion Recognition in Conversations.
- **HCI:** human-computer interaction

## Comparación [1]

Esa comparación estará enfocada entre los documentos "A Review on Speech Emotion Recognition Using Deep Learning and Attention Mechanism" (Ref. 2) y "Af-CAN: Multimodal emotion recognition method based on situational attention mechanism" (Ref. 4).

Este último propone un nuevo método; Af-CAN (Attention-based Fusion for Context-aware Networks) para el reconocimiento de emociones en diferentes tipos de datos (texto/video/audio). La integración de estos diferentes tipos de datos mejora la precisión en el reconocimiento de emociones.

Los dos documentos destacan la importancia de los mecanismos de atención en el reconocimiento de emociones, aunque desde perspectivas diferentes. Ref.4 propone una solución innovadora para el reconocimiento multimodal (es decir, de diferentes tipos de datos), mientras que Ref.2 muestra una visión general de técnicas en el reconocimiento emocional basado en el habla.

Las diferencias más destacables entre los dos documentos son:

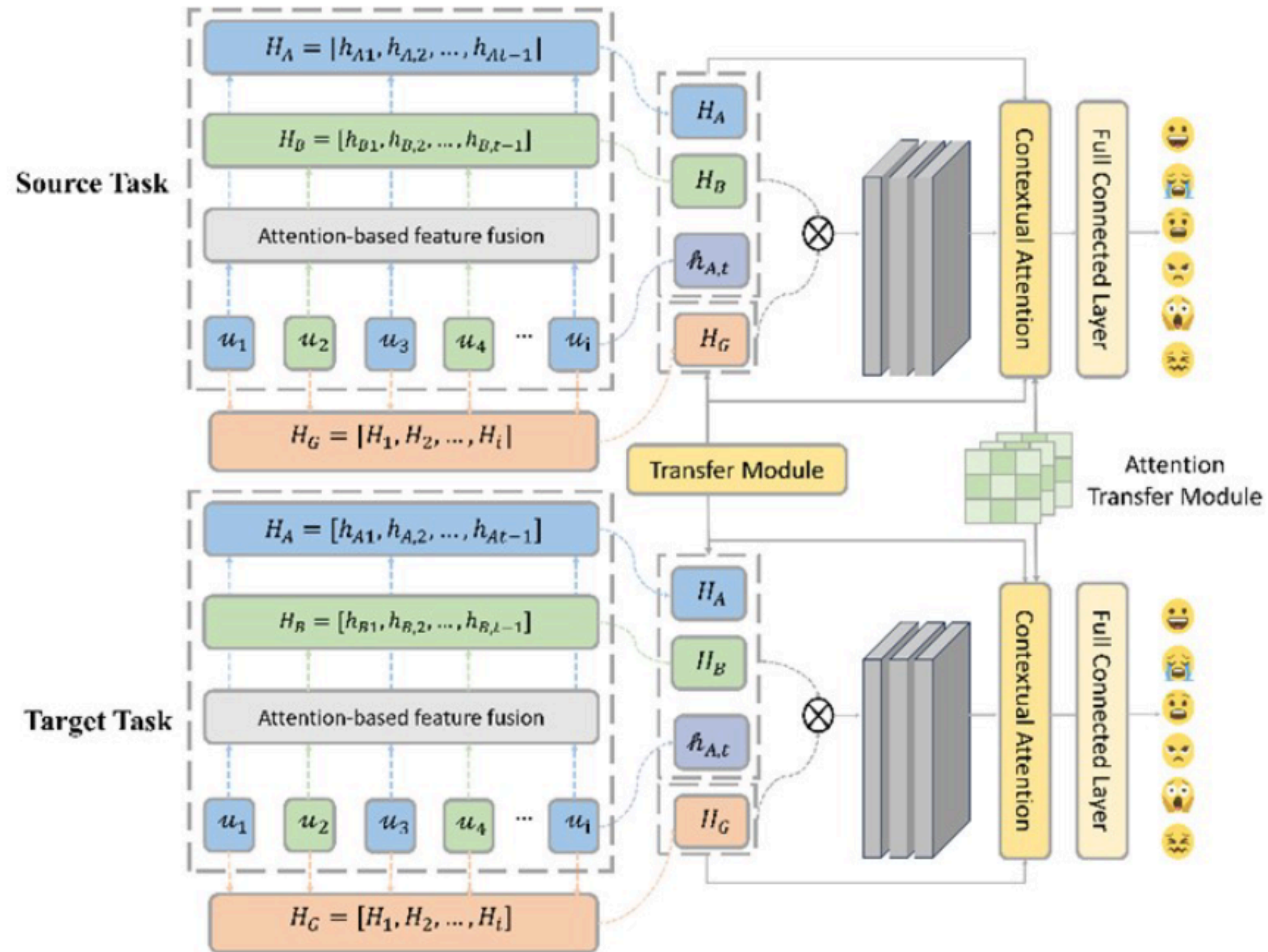
Week 9

Research Stay

# Comparación [2]

[Ref 2] "A Review on Speech Emotion Recognition Using Deep Learning and Attention Mechanism"	[Ref 4] "Af-CAN: Multimodal emotion recognition method based on situational attention mechanism"
Utiliza únicamente datos de audio	Usa diferentes tipos de datos (audio, texto y video)
Busca mejorar el reconocimiento de emociones en el habla	Mejora la precisión del reconocimiento de emociones usando un mecanismo de atención situacional
No propone como tal una arquitectura. Analiza diferentes métodos (como RNN y CNN) para el reconocimiento de emociones y mecanismos de atención.	Su arquitectura cuenta con módulos para cada modalidad (tipo de dato) y un mecanismo de atención situacional que ajusta la atención según el contexto.
Proporciona una visión general de las técnicas actuales y sus respectivos resultados en tareas de reconocimiento de emociones. Menciona desafíos del campo, pero no presenta un nuevo método o modelo .	Presentan los resultados del modelo propuesto, que demuestra mejoras significativas en la precisión del reconocimiento de emociones al integrar diferentes tipos de datos. El mecanismo de atención situacional es una contribución importante.
Su aplicación puede limitarse solo a entornos donde el audio es la fuente de información	Puede ser aplicado en una variedad más amplia debido a la integración de diferentes tipos de datos





**FIGURE 2.** The Af-CAN module is detailed with three primary components: A) feature extraction and fusion module, B) emotion classification network driven by contextual attention mechanisms, and C) transfer learning module for emotion recognition across different corpora.

# Referencias

- [1] Attention is all you need
- [2] A Review on Speech Emotion Recognition Using Deep Learning and Attention Mechanism
- [3] Attention mechanisms in computer vision: A survey.
- [4] Af-CAN: Multimodal emotion recognition method based on situational attention mechanism