RESEARCH PAPER

AR-EAMH: An Emotion-Aware Mental Health Prediction Framework Using Self-Attention-Based Transformers for Adaptive Rehabilitation

P Sudhandradevi[1], V Bhuvaneswari[2*]
[1]Ph.D. Research Scholar, Department of Computer Applications, Bharathiar University, Tamil Nadu, India.
Email: psudhandradevi@gmail.com
[2]Professor, Department of Computer Applications, Bharathiar University, Tamil Nadu, India.
Email: bhuvana_v@buc.edu.in
***Corresponding Author: V Bhuvaneswari**
Professor, Department of Computer Applications, Bharathiar University, Tamil Nadu, India
Email: bhuvana_v@buc.edu.in

ABSTRACT
Artificial Intelligence (AI)-driven affect recognition is transforming mental healthcare by enabling real-time monitoring of emotional states such as depression and allowing intelligent systems to adapt dynamically on users' emotional responses. Especially attention mechanism, which revolutionized Natural Language Processing (NLP) by choosing long range dependencies and higher-order affect patterns such as sarcasm, negation, and contextual sentiment changes, providing better scalability and performance compared to existing sequence models such as Recurrent Neural Networks (RNNs). GANs produce synthetic data, for example, emotional discussions and locomotion patterns, to address dataset availability and providing personalized therapy exercises. Transformers (GPT-family), which are powered by multi-head self-attention, enable conversational agents to empathize, generate rehabilitation programs, and provide real-time adaptive feedback. In this paper, a personalized emotion recognition model to develop Emotion-Aware AI for Mental Health (EAMH) and by integrating the theories of Ekman and Plutchik's to enhance sentiment analysis. The paper processed in the following pipeline: (1) Data acquisition and Text preprocessing, (2) BERT-based feature extraction, (3) contextual understanding, (4) semantic mapping, (5) self-attention to preserve emotional nuances, and (6) multi-head attention to preserve diverse emotional representations. Through these mechanisms, it can accurately differentiate between depression-related emotions states. The findings validate the efficacy of attention-based AI in identifying complex emotional patterns, leading to adaptive AI interventions and customized support systems in digital healthcare.
Keywords: Contextual Relationships; Deep Learning; Emotion Recognition; Multi-Head Attention; Natural Language Processing; Self-Attention Mechanism; Transformers.
How to cite this article: Sudhandradevi P, Bhuvaneswari V. AR-EAMH: An Emotion-Aware Mental Health Prediction Framework Using Self-Attention-Based Transformers for Adaptive Rehabilitation. Int J Drug Deliv Technol. 2026;16(59s): 454-466. DOI: 10.25258/ijddt.16.59s.48
Source of support: Nil
Conflict of interest: None

1. INTRODUCTION

Artificial Intelligence (AI) systems growing responsibility to demonstrate Emotional Intelligence (EI), as an ability to discern, interpret, and provide reactions of feelings in speech, writing, and facial expressions (Goleman, 1995). The lack of identifying the depression signs is overlooked in conventional testing where AI-powered emotion recognition models to detect subtle cues in language, tone, and sentimens as earlier. The need of AI driven depression detection is expressed through face-to-face communication in online platforms to improve early identification of patient condition to enable scaleable data-driven mental health care systems. The technology equips health practitioners with actionable information, closing clinical intervention to digital expression, ensuring timely and personalized intervention for people at risk. In recent years multi-modal fusion blends the text, speech, and facial expression to identifies emotion classification (Poria et al., 2020) to assess mental health depression and anxiety from unstructured data (Schuller et al., 2021). To address this issue, convolutional neural networks (CNNs) identify text-based emotion recognition (Kim, 2014) to extract sentiment patterns and local dependencies and in speech emotion recognition it extracts the spectrogram features (Tzirakis et al., 2017) with long-range dependencies in sequential data processing.

Transformers (Vaswani et al., 2017) and self-attention mechanism revolutionized NLP to handle a full sequence in parallel to capture long-range dependencies. Attention mechanisms is integrated with RNN and LSTM dynamically with the relevance of different words or speech frames to enhance sentiment and emotion classification tasks (Bahdanau, et al., 2015). Models that combined CNNs, LSTMs, and attention mechanisms to optimize emotion classification performance by leveraging capabilities from many architectures. However, this technique is relied on sequential processing to increase the scalability. Self-attention mechanisms, BERT and GPT brought about significant improvements in emotion detection by dynamically focusing on contextual cues in a sentence. These advances have subsequently given rise to subsequent-of-art AI systems that are able to effectively diagnose and interpret human feelings in a sequence of multi-modalities such as text, voice, and face expressions. Self-attention-based AI systems have been promising to detect mental disorders such as depression and anxiety by analyzing patient conversations, social media messages, and clinical documentation (Shen et al., 2020). Such models may be able to capture subtle cues of emotion that may indicate psychological distress and permit early intervention and personalized mental health care.

Generative AI Integration for Rehabilitation

Generative AI techniques empower adaptive rehabilitation program GAN to synthesize artificial rehabilitation data (e.g., emotional dialogues, movement patterns) to exceed dataset sparsity and generate individualized exercise for therapy. VAE to generate adaptive rehabilitation material model for patient-specific recovery paths. Transformers (GPT-family) to enable empathetic conversational agents and generate personalized rehabilitation plans to provide real-time adaptive feedback of patients. This study presents a novel attention-based framework Emotion-Aware AI for Mental Health (EAMH) for emotion recognition in mental health contexts. The major contributions are:

- Theory-Driven Emotion Mapping: EAMH combines Ekman's and Plutchik's psychological theories to extrapolate GoEmotions labels.
- Extended Attention Mechanism: A fine-tuned self-attention pipeline, with multi-head attention, identifies subtle cues such as sarcasm and contextual change of sentiment commonly overlooked by standard models.
- Psychiatric Health Priority: Differing from standard emotion classifiers, EAMH is specifically designed for depression-related emotion recognition, filling a key void in NLP.

The following section recommends the current emotion-aware recognition models, in order to give some basis literature on which to build the rest of the sections of this chapter.

2. Literature Review

Devlin et al. (2019) presented BERT a pre-trained bidirectional transformer model which set new performance benchmarks on a wide range of NLP tasks. Mickel Hoang et al. (2019) transferred BERT to out-of-domain Aspect-Based Sentiment Analysis (ABSA) with higher accuracy than the current models. In ABSA Akbar Karimi et al. (2020) added parallel and hierarchical aggregation modules to enhance the performance of BERT in aspect extraction and sentiment classification. Hu Xu et al. (2020) analyzed the attention heads in BERT, which indicated that not many of them capture aspect and opinion words and most retain domain-specific semantics. Dosovitskiy et al. (2020) introduced the Vision Transformer (ViT) that inherently processes sequences of image patches, achieving convincing performance on image classification. In speech emotion recognition (SER), Anish Nediyanchath et al. (2020) proposed a multi-head attention model with positional embeddings and gender recognition as a support task and achieved 76.4% accuracy on the IEMOCAP dataset. Ping Ji & Xiaohong Liu (2020) extended BERT with multi-head attention and positional weights to perform on datasets.

Siriwardhana et al. (2020) discussed self-supervised aspects of pre-trained text, audio, and vision models and suggested a Transformer-based fusion scheme. Such models such as Wu et al. (2023) have all been persistently state-of-the art on IEMOCAP, CMU-MOSEI, and MELD datasets, emphasizing how good multi-modal fusion is instrumental in enhancing emotion recognition. Mingke Xu et al. (2021) developed a head fusion method over multi-head attention with 76.18% weighted and 76.36% unweighted accuracy in IEMOCAP. They also conducted noise injection experiments to guarantee model robustness. These results indicate how multi-head attention aids SER with enhanced accuracy and model robustness under varying conditions. In facial expression recognition (FER), Yu & Bai (2021) proposed a visual self-attention network to integrate local image features and cancel out external interference. Wen et al. (2021) proposed a two-stream global-guided attention network (TGGAN) to describe the interaction between global and local facial features. Pecoraro et al. (2021) suggested local multi-head self-attention module to enhance FER2013 performance. Zhou et al. (2022) proposed RSACNN region self-attention with CNN that leverage expression-salient facial textures with robust performance on CK+, FER2013, and RAF-DB. Shen & Xu (2023) proposed a multi-head attention module integrated with ResNet18 framework on different datasets. Song J (2023) used Vision Transformers for FER, with improved recognition of local patterns using new architectural adjustments. To mitigate emotional asynchrony and misalignment in modality, a multi-granularity attention mechanism was put forth by Fan et al. (2023).

Dutta & Ganapathy (2023) integrated recurrent and co-attention networks within a hierarchical cross-attention architecture for audio-text emotion recognition. Yue Xie, et al., (2023) upgraded LSTM models using multi-head attention with time and feature dimensions, borrowing attention mechanisms to sequential modeling. Zhang et al., (2023) proposed DualGATs, a dual-graph attention network combines conversational context, modality interaction, and multi-task objectives in sentiment analysis and emotion recognition. Xiao et al. (2024) also enhanced VGG16 with SGE attention modules and feature fusion, achieving over 90% accuracy on typical facial expression datasets. Xinxin Jiao et al. (2024) presented MFHCA, which fuses multi-spatial fusion and hierarchical cooperative attention, and showed better performance on IEMOCAP. The mechanism of attention has transformed emotion recognition by allowing models to attend to salient features without noise. Self-attention mechanisms, as in Transformers and BERT-based models, enhance sentiment analysis through the ability to capture long-distance dependencies and context relationships. Multi-head attention also improves feature representation and is a crucial building block in NLP, speech analysis, and facial expression.

3. FRAMEWORK AND METHODOLOGY: EAMH

Social media platforms generate vast amounts of text data were users express emotions openly. Identifying at-risk individuals, analyzing public sentiment trends, and enhancing AI chatbot interactions require fine-grained emotion detection. The Emotion-Aware AI for Mental Health (EAMH) framework is extended to explicitly incorporate generative AI techniques for rehabilitative applications. Beyond detecting and classifying emotions, EAMH leverages Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), Transformer-based architectures, and Diffusion Models to create adaptive, personalized, and scalable rehabilitation solutions. The methodology is divided into four phases is given in Figure 1.

1. Data Acquisition and Text Pre-Processing
2. Feature Extraction & Contextual Understanding
3. Semantic Mapping (Ekman & Plutchik Models)
4. Self-Attention and Multi-Head Attention-Based Emotion Learning

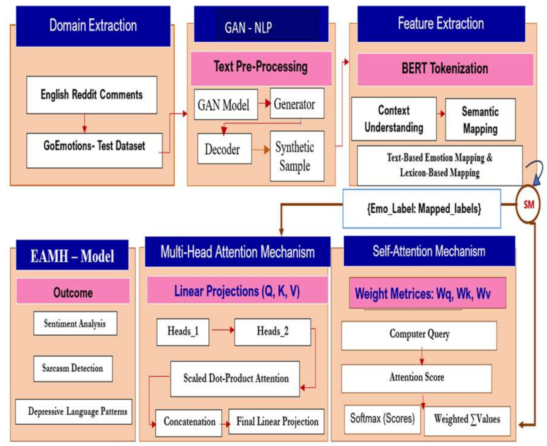

Figure 1: The Emotion-Aware AI for Mental Health (EAMH) framework

PHASE I: DATA ACQUISITION & TEXT PRE-PROCESSING

The GoEmotions widely used dataset in Natural Language Processing (NLP) and Emotion Recognition due to its rich, multi-label emotion annotations. A comprehensive dataset, after pre-process containing 63,685 English Reddit comments, annotated with 27 emotion categories, including joy, sadness, anger, and more. This dataset offers fine-grained emotion labels suitable for training models in nuanced emotion detection is given in Table 1.

Table 1: Fine-Grained Emotion Labels

Category	Emotions
Positive Emotions	Admiration, Joy, Gratitude, Excitement, Love, Optimism, Pride, Relief
Negative Emotions	Anger, Disappointment, Fear, Sadness, Disgust, Grief, Embarrassment, Disapproval, Remorse, Annoyance
Ambiguous/Mixed Emotions	Curiosity, Realization, Surprise, Confusion, Desire, Nervousness, Caring
Neutral	Represents cases where no strong emotion is detected.

NLP IN MENTAL HEALTH REHABILITATION

Natural Language Processing (NLP) forms the core of AI-driven emotion recognition systems, enabling machines to understand, interpret, and analyze human emotions from textual data. With the increasing digitalization of human communication through chatbots, social media, and virtual therapy sessions NLP-based emotion recognition is transforming mental health monitoring, sentiment analysis, and personalized AI interactions. In mental health application areas, AI-driven NLP models review patient conversations, diaries, or sessions to detect emotional distress such as depression, anxiety, or frustration. This facilitates early detection, intervention, and continued monitoring. NLP extends simple polarity (positive, negative, neutral) sentiment analysis by detecting subtle emotions such as anger, joy, sorrow, and sarcasm. Whereas standard NLP pipelines utilize real-world data, shortage and imbalance are still some of the most critical challenges to mental health usage. Generative Adversarial Networks (GANs) are now incorporated into NLP pipelines for rehabilitation to overcome this.

- Synthetic Data Generation: GANs generate realistic copies of depressive or anxiety-related statements without exposing sensitive patient data.
- Data balancing: By filling gaps in underrepresented emotional categories, GANs enable NLP models to be trained on a well-balanced dataset.
- Robustness: Exposure to diverse synthetic samples improves the model's ability to handle unseen or rare emotional expressions.
- Personalization: Conditional GANs can generate statements reflecting different severity levels or therapy contexts, enabling personalized rehabilitation plans.

```
#1 GANs for NLP-based Mental Health
Input: Real depressive text dataset
1. Preprocess:
  - Clean & tokenize text
  - Convert to embeddings
2. Define GAN:
  - Generator G(z) → synthetic embeddings
  - Discriminator D(x) → real vs fake
3. Train:
  For epochs:
    - Sample real_embeddings
    - Generate fake_embeddings = G(z)
    - Update D to classify real/fake
    - Update G to fool D
4. Generate:
  - Use G(z) to create new depressive embeddings
  - Decode embeddings back to text
5. Augment:
  - Combine synthetic + real data
  - Train final emotion/depression classifier
Output: Balanced dataset & robust NLP model
```

PHASE II: FEATURE EXTRACTION & CONTEXTUAL UNDERSTANDING

BERT tokenizer is a critical building block in emotion recognition applications with NLP, primarily used in processing datasets. In contrast with standard tokenization techniques, BERT utilizes WordPiece tokenization, which splits words into semantic subwords, enabling it to perform very well with rare or out-of-vocabulary words that often feature in emotional sentiments. The technique is such that even complicated words like "unhappiness" are divided into ["un", "##happiness"], keeping their semantic sense intact. BERT enable models to differentiate between comparable-looking sentences with contrastive sentiment, i.e., "I am happy" vs. "I am not happy." As BERT is pre-trained on a large corpus, including Books corpus and Wikipedia, it picks up subtle feelings across various linguistic styles, rendering it very versatile for emotion analysis in real-world scenarios. Further, BERT is good at recognizing sarcasm and sentiment change-around, which is common in the GoEmotions dataset. For example, a sentence such as "I was happy, but now I am disappointed" takes a change of sentiment that standard tokenizes cannot convey. By leveraging Self-Attention and Multi-Head Attention mechanisms, BERT ensures that contextual dependencies between words are properly modeled, enabling a more accurate representation of emotions. This makes BERT's tokenizer a fundamental choice when working with transformer-based emotion recognition models.

```
#2 Feature Extraction
BEGIN
# Step 1: Load Dataset
LOAD dataset from "goemotions.csv"
# Step 2: Remove Duplicates
[REMOVE duplicate rows from dataset
# Step 3: Reformat Dataset
EXTRACT "text", "id", and "emotion labels" columns
CONVERT emotion labels into multi-label format (one-hot encoding)
# Step 4: Load BERT Tokenizer
IMPORT BERT tokenizer from transformers library
INITIALIZE tokenizer with "bert-base-uncased"
# Step 5: Tokenization
FOR each text sample in dataset:
   TOKENIZE text using BERT Tokenizer
   CONVERT tokens into token IDs
   TRUNCATE or PAD sequences to max length
   STORE processed tokens
# Step 6: Convert Data into Tensor Format
CONVERT tokenized text into tensors
CONVERT emotion labels into tensor format
END
```

These embeddings not only serve classification tasks but also condition generative models (e.g., conditioning a VAE or diffusion model on detected emotional states to simulate patient behavior or adapt therapeutic content).

Context Understanding

Emotion recognition in text is challenging due to linguistic complexities like sarcasm, negation, and context-dependent sentiment changes. Sarcasm Detection: NLP models must recognize cases where a phrase's intended meaning is opposite to its literal interpretation. Negation Handling: Words like *"not", "never", "hardly,"* can reverse the meaning of a sentence. Context-Based Emotion Shifts: Some sentences may start positively but conclude negatively.

```
#3 Context Understanding
BEGIN
  IMPORT sentiment analysis pipeline from transformers
  # Step 1:  Load the pre-trained sentiment classifier
  LOAD classifier using sentiment-analysis model
  DEFINE sentences AS [   # Define input text samples
      "I'm not feeling great today.",
      "I just love when my internet disconnects in the middle of work!",
      "I was happy, but now I'm really disappointed."   ]
  # Step 2: Perform sentiment prediction
  results ← APPLY classifier ON sentences

  # Step 3:  Iterate through results and print output
  FOR each sentence, result IN sentences, results DO
      EXTRACT sentiment label and confidence score
      PRINT "Text: [sentence] | Sentiment: [result.label] | Confidence: [result.score]"   END FOR
END
```

Phase III: Semantic Mapping (Ekman & Plutchik Models)

Semantic mapping in emotion recognition involves classifying extracted features, such as text, speech, or facial expressions, into structured emotional categories based on their meaning and context. This process is essential for accurately mapping emotions to predefined models like Ekman's model and Robert Plutchik's model in Table 2. Paul Ekman's model (1992) is widely accepted because it identifies six fundamental emotions (anger, happy, sad, surprise, disgust, and fear) and that are cross-culturally recognized through facial expressions. These emotions form the core of human affective responses and are used in many emotion detection systems, including facial expression analysis and sentiment classification. Robert Plutchik's model (1980) extends Ekman's model by incorporating eight primary emotions (sadness, joy, trust, fear, surprise, anger, disgust, and anticipation), and further classifies them into intensities (e.g., annoyance → anger → rage). The model also supports emotion blending, enabling sophisticated emotional states such as optimism (joy + anticipation) and love (joy + trust). Mapped emotions are control signals for generative AI models. For example, recognizing "anxiety" can be used to steer a generative conversational agent (Transformer-based) to generate soothing responses, or a GAN-based system to mimic safety exposure therapy situations.

Table 2: Semantic Mapping: Use Case

Detected Emotion (GoEmotions)	**Mapped Emotion (Custom Domain-Specific Labels)**	**Relevance**
Disappointment	Frustration	Detecting early signs of distress in mental health assessments
Nervousness	Anxiety	Identifying anxiety-related behaviors in conversations
Grief	Depression	Monitoring severe emotional distress and crisis situations
Remorse	Guilt	Understanding emotional responses in therapy & self-reflection
Annoyance	Irritation	Detecting user frustration in customer support interactions
Embarrassment	Social Anxiety	Identifying cyberbullying victims or self-esteem issues
Disapproval	Judgment / Rejection	Tracking online hate speech and

		negative sentiment
Optimism	Hope	Analyzing positive sentiment trends in social media
Admiration	Inspiration	Understanding motivation and leadership-driven emotions
Excitement	Motivation	Tracking engagement in online learning & workspaces
Love	Emotional Bonding	Detecting relationship-driven emotions in conversations
Caring	Empathy	Identifying supportiveness in mental health communities
Surprise	Shock / Alert	Recognizing unexpected reactions in crisis management

Semantic mapping enables AI models to understand words, phrases, tone, and context to decide on the most applicable emotional state. By linking your emotions to key topics, this customized system monitors your feelings in real-time, allowing AI to understand them better and provide useful information. Explainable (XAI) is improved along with mapping emotions effectively, enhancing AI-powered responses in chatbots and enabling real-time emotion tracking in mental health use cases. This systematic method ensures emotion classification is meaningful and contextually valid, allowing AI to deliver personalized and adaptive interactions.

#4 Text-Based Emotion Mapping & Lexicon-Based Mapping

```
Step 1:
  Function Preprocess_Text(text):
   Convert text to lowercase
   Remove punctuation and special characters
   Tokenize text into words
   Return list of words
Step 2:
Function Load_Emotion_Lexicon():
   Define a dictionary of words mapped to emotions
   Return the lexicon
Step 3:
Function Map_Emotions(text):
   words ← Preprocess_Text(text)
   emotion_counts ← Empty dictionary

   For each word in words:
      If word exists in Emotion_Lexicon:
         Increase count of corresponding emotion in emotion_counts
       If emotion_counts is not empty:
          Return emotion with highest count
   Else:
      Return "Neutral"
# Step 4: Example Usage
     sentences ← ["I am so happy and excited today!", "I feel sad and hopeless.", "I am very angry and furious about this!", "I was shocked by the unexpected news.", "This is absolutely disgusting!", "I am feeling a little nervous."]
Step 5: For each sentence in sentences:
            Emotion ← Map_Emotions(sentence)
            Print "Text:", sentence, "| Emotion:", Emotion
```

PHASE IV: SELF-ATTENTION AND MULTI-HEAD ATTENTION-BASED EMOTION LEARNING

Traditional models face challenges in handling long-term dependencies as sequences grow; earlier words tend to lose significance of vanishing gradients in recurrent architectures. For example, in the sentence "I was excited, but now I'm disappointed," traditional models treat "excited" and "disappointed" as independent words, failing to recognize the emotional contrast. As opposed to RNNs, which process sequences sequentially, Transformers can process whole input sequences in parallel, solving problems such as slow training, vanishing gradients, and inability to capture long-range dependencies. In attention model types, we are interested in self-attention and multi-head attention. Self-attention eliminates this constraint by dynamically assigning varying levels of importance weights to words depending on their context. Rather than processing words sequentially as in LSTMs, Self-Attention simultaneously considers all words so that highly emotionally charged words are given greater importance. For example, in the given sentence, "excited" and "disappointed" would be given more importance than filler words such as "but" or "now." This process allows the model to pay attention to important emotional words while rejecting less important ones. In addition, self-attention successfully catches long-range dependencies so that it can grasp sentiment changes and contextual subtleties without losing information. Through the examination of a complete sentence, it guarantees that key emotional signals, even distant ones, are all factored into the ultimate sentiment label. Multi-head attention subsequently amplifies this by being able to capture several contextual relationships at a time. As Transformers do not have sequential information, the word order is preserved using positional encoding. Feed-forward layers also transform the extracted features, and residual connections and layer normalization make training more stable. Transformers have emerged as the unsung pillars of contemporary AI applications, fueling models such as BERT to understand context, GPT to generate text, and T5 to summarize text and translate it. They are extensively used in sentiment analysis, emotion detection, speech-to-text translation, and multimodal AI tasks and have been a breakthrough in deep learning and NLP.

Self-attention Mechanism

Self-attention mechanism enables a model to weigh different parts of an input sequence when making predictions. In traditional models RNNs process step-by-step, whereas in self-attention processes all elements in parallel and capturing long-term dependencies efficiently. Self-attention dynamically adjusts how words (or speech frames) relate to each other, improving tasks like emotion recognition. Self-attention operates using three key matrices derived from input embeddings:

Working of Self-Attention Mechanism

For an input sequence as a matrix X will generate, is the focus of the current token as query (Q), Key (K) as the comparison reference, and stores the actual token contents as Value (V). After the model decides how much attention to assign to every word with value vectors, to create a weighted representation of the context.

Step 1: Calculate Query, Key, and Value Matrices

$$Q = XW_Q,\quad K=XW_K,\quad V=XW_V \tag{1}$$

where W_Q, W_K, and W_V are learned weight matrices.

Step 2: Compute Attention Scores

The attention score is calculated using a dot product to find similarity between queries and keys,

$$\text{Score}(Q, K) = QK^T \tag{2}$$

To prevent large values, scale by $\sqrt{d_k}$, where **d_k** is the dimensionality of K,

$$\text{Scaled Attention}(Q, K) = \frac{QK^T}{\sqrt{d_k}} \tag{3}$$

Using Softmax guarantees that the scaled scores and then undergo a softmax function to normalize them into a probability distribution whose attention weights add up to 1. The probability values that determine the priority of the probability distribution to be inserted in the sequence.

$$\alpha = \text{softmax}(\text{Scaled Attention}) \tag{4}$$

Step 3: Compute Weighted Sum of Values

The output vectors' results for every one representing a word in the sequence enhance the entire sequence to enhance the contextual information from the entire sequence. Every token sums up others' information through attention scores:

$$\text{Attention}(Q, K, V) = \alpha V \tag{5}$$

Multi-head Attention

Rather than relying on one attention function, Transformers employ multi-head attention, and the model learns several relationships among words or frames. Different linguistic or emotional relationships are captured by each head. In parallel, the model calculates several attention scores rather than one attention score. Example of "emotion recognition". With the application of multiple attention heads, the model can capture a more comprehensive set of dependencies in the input sequence. For instance, one head can pay attention to the structure of the complete sentence, whereas another can attend to individual details. By integrating all these varied views, Multi-Head Attention serves a richer insight into the input, similar to how humans perceive various aspects of information together at once. Multi-head attention involves the replication of the self-attention task various times, for various linear projection input is illustrated in Figure 2. This enables the model to attend various aspects of the self-attention task simultaneously.

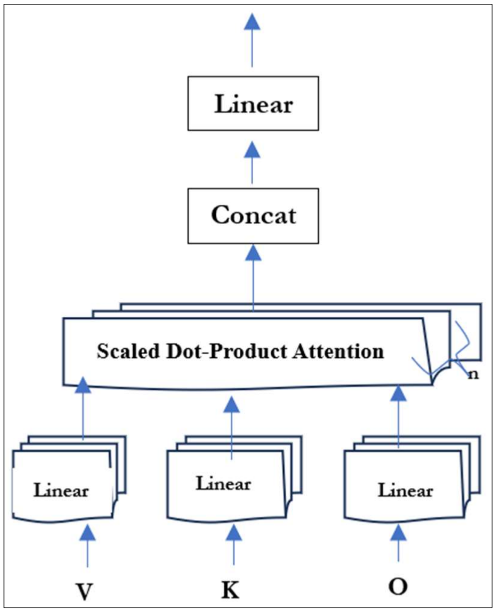

Figure 2: Multi-Head Attention

Step 1: Linear Projection

For every head in attention, the input sequence is projected linearly into queries(Q), keys(K) and values(V) are projected linearly with the unique learned weights matrices. They form various versions of heads as depicted above Figure 2.

$$Q_h = XW_{Qh},\ K_h = XW_{Kh},\ V_h = XW_{Vh} \tag{6}$$

Here h indicates the different heads, where W_{Qh}, W_{Kh} and W_{Vh} are learned weight matrices for the h-th head.

Step 2: Scaled Dot-Product Attention

The model computes attention scores, scales them, and uses softmax function and produces context-sensitive outputs for every set of these metrices Q, K, and V go through scaled dot-product independently.

$$\text{Attention}_h\ (Q_h, K_h, V_h) = \text{softmax}\left(\frac{Q_h K_h^T}{\sqrt{d_k}}\right) V_h \tag{7}$$

Step 3: Concatenation

Once all of them have been processed individually through the attention heads, the output of all heads(h) gets concatenated together. This step unites the various views that each head has concentrated on.

$$\text{Concatenated\_Output} = \text{Concat}\ (\text{Attention}_1, \text{Attention}_2 \ldots\ldots, \text{Attention}_h) \tag{8}$$

Step 4: Final Linear Projection

The concatenated output is fed into a second linear projection with a weight matrix (WO) in order to generate the final output. In this step, all the various points of view are combined into one unified representation:

$$\text{MultiHead}\ (Q, K, V) = \text{Concatenated\_Output} \times W_O \tag{9}$$

#5. Multi-Head Attention
BEGIN
Input: Word embeddings (X)
Define number of heads (h)
Initialize weight matrices for each head (Wq, Wk, Wv)
Compute Q, K, V for each head:
for each head h:
Q_h = X * Wq_h
K_h = X * Wk_h
V_h = X * Wv_h
Compute self-attention for each head:
attention_h = softmax((Q_h * K_h^T) / sqrt(d_k)) * V_h
Concatenate outputs from all heads
Apply final transformation using W_o
Return
END

4. RESULTS AND DISCUSSION

Use Case: Real-time Emotion Detection in Social Media Posts

Social media websites generate immense quantities of user-generated content with a broad spectrum of emotions. Real-time emotion detection is important for detecting mental health threats, cyberbullying, distress signals, and nascent public opinion trends. Most traditional sentiment analysis models are not adept at fine-grained emotion classification and tend to look subtle changes, sarcasm, and mixed sentiments. Sophisticated NLP methods, including multi-head attention and self-attention, enhance context awareness by dynamically weighing words and modeling long-distance dependencies. On data such as GoEmotions, models can label emotions more accurately, separating subtle emotions such as disappointment, pride, or curiosity. Live emotion monitoring is an advantage to social media websites, mental health clinicians, companies, and policymakers, allowing for preemptive responses to crises, targeted content suggestions, and improved user engagement. By utilizing Transformer-based models, this technology offers richer emotional intelligence, ensuring social media is safer, more responsive, and emotionally intelligent.

Step 1: Data Pre-Processing

The dataset comprises Reddit posts, user IDs, and 27 emotion labels and thus forms a rich resource for emotion classification in social media text. What these annotations reveal are insights into how emotions are labeled so that model training becomes more effective. For the dataset not to become biased and diversified, duplicate posts are eliminated. Duplicate entries bias model predictions towards overrepresented emotions. By removing duplicates, the dataset retains its representativeness and improves model reliability. To streamline the processing pipeline, unnecessary metadata (such as subreddit names and timestamps) is removed, retaining only the text, post ID, and emotion labels is shown in Figure 3.

	text	emo_labels
0	worst ending ever! i won't spoil it but this o...	'disappointment'
1	happy cake day u sneakpeekbot!	'excitement'
2	was he rejected because of his methodology or ...	'confusion'
3	thanks, i agree	'approval', 'gratitude'
4	why would you doubt it dumbass?	'anger', 'disgust'
...	...	...
71216	tell her youre sleeping with someone else you...	'approval'
71217	you got banned for participating in a brigade	'annoyance', 'approval'
71219	well, i'm glad you're out of all that now how...	'joy'
71221	well when you ve imported about a gazillion of...	'caring'
71223	the fda has plenty to criticize but like here...	'anger'

Figure 3. Dataset Structure: Goemotions

The dataset consists of 63685 instance and 4 columns. Since machine learning models cannot process raw text, each post is converted into numerical tokens using BERT's tokenizer. Tokenization breaks down text into smaller components that the model can interpret while preserving its semantic meaning. This step ensures that contextual relationships between words remain intact, allowing the model to understand emotional nuances in the text. BERT's tokenizer also effectively handles slang, abbreviations, and informal language, which are common in social media posts. A single post can express multiple emotions simultaneously. Thus, emotion labels are converted into a binary multi-label format, where each emotion is represented as 1 (present) or 0 (absent). By representing emotions in this format (Figure 4), the model can detect combinations of emotions, leading to a more accurate and nuanced classification system

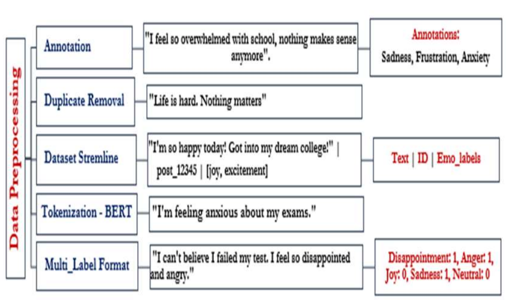

Figure 4: Data Pre-Processing

In NLP-based mental health applications, GANs can mitigate data scarcity and imbalance for underrepresented categories like depressive statements. Real text samples are embedded into a numerical space using models such as SentenceTransformer, capturing their semantic meaning.

```
Synthetic Sample 1:
  ~ Everything feels meaningless and exhausting.   (similarity: 0.30)
  ~ I am hopeless and tired of trying.   (similarity: 0.28)

Synthetic Sample 2:
  ~ Everything feels meaningless and exhausting.   (similarity: 0.30)
  ~ I am hopeless and tired of trying.   (similarity: 0.30)

Synthetic Sample 3:
  ~ I am hopeless and tired of trying.   (similarity: 0.29)
  ~ I feel lost. No motivation. Nothing excites me.   (similarity: 0.21)

Synthetic Sample 4:
  ~ Everything feels meaningless and exhausting.   (similarity: 0.30)
  ~ I feel lost. No motivation. Nothing excites me.   (similarity: 0.29)

Synthetic Sample 5:
  ~ I feel empty and drained every day.   (similarity: 0.31)
  ~ I am hopeless and tired of trying.   (similarity: 0.29)
```

Figure 5: Top 5 Synthetic Text samples generated with their Corresponding Similarity Scores

The GAN trains in this space, with the generator producing synthetic embeddings and the discriminator distinguishing real from synthetic samples, resulting in diverse, realistic depressive text representations in Figure 5. Visualization methods such as t-SNE validate the intersection between real and generated embeddings, confirming GAN-based data augmentation for better classification robustness.

STEP 2: CONTEXT UNDERSTANDING

Traditional sentiment analysis, to classify a text as positive, negative, or neutral, context comprehension goes beyond how words and phrases combine to create meaning. In the recognition of emotion, because it ensures that emotions are accurately understood in view of text nuances, sentiment change, and intensity nuances. Sarcasm recognition, in which a phrase has a negative connotation even though it employs a positive word. Another critical ingredient is sentiment shifts, which track changes in emotions within a sentence and signal a shift from one emotion to another. Emotion intensity analysis quantifies how intensely an emotion is being expressed. The labeled sentiment, sentiment score, and sarcasm tags, saved in a structured format for potential further analysis and incorporation into the overall emotion recognition model, are depicted in Figure 6. Utilizing both sentiment and sarcasm detection enables the AI model to acquire a more accurate interpretation of a patient's emotional state, minimizing misunderstandings and the precision of depression detection.

	text	emo_labels	sentiment_label	sentiment_score	sarcasm
0	worst ending ever! i won't spoil it but this o...	'disappointment'	NEGATIVE	0.999641	Not Sarcastic
1	happy cake day u sneakpeekbot!	'excitement'	POSITIVE	0.991722	Not Sarcastic
2	was he rejected because of his methodology or ...	'confusion'	NEGATIVE	0.998824	Not Sarcastic
3	thanks, i agree	'approval', 'gratitude'	POSITIVE	0.999832	Not Sarcastic
4	why would you doubt it dumbass?	'anger', 'disgust'	NEGATIVE	0.997942	Not Sarcastic
...	...	...	...	...	...
71216	tell her youre sleeping with someone else you...	'approval'	NEGATIVE	0.998927	Not Sarcastic
71217	you got banned for participating in a brigade	'annoyance', 'approval'	NEGATIVE	0.997350	Not Sarcastic
71219	well, i'm glad you're out of all that now how...	'joy'	NEGATIVE	0.980157	Not Sarcastic
71221	well when you ve imported about a gazillion of...	'caring'	POSITIVE	0.787407	Not Sarcastic
71223	the fda has plenty to criticize but like here...	'anger'	NEGATIVE	0.999036	Not Sarcastic

Figure 6. Sarcasm Detection

STEP 3: SEMANTIC MAPPING: INDENTIFYING DEPRESSIVE LANGUAGE PATTERNS

Semantic mapping, where detected emotions are systematically categorized based on established psychological models. This process ensures that the detected emotions align with recognized emotional theories, enhancing interpretability and clinical relevance. In this application, GoEmotions dataset emotions are mapped to three models: Ekman's Basic Emotions Model, Plutchik's Wheel of Emotions, and a Custom Mental Health Mapping, is depicted in Figure 7. These mapped emotions are then utilized for real-time emotion tracking, aiding the detection of signs of depression, anxiety, or stress. The formal representation also aids explainability and interpretability, essential for adopting AI in mental health diagnostics.

```
Text: worst ending ever! i won't spoil it but this one sucked worse than the avengers last movie!
Emo Labels: 'disappointment'
Mapping Col: Sadness
--------------------
Text: happy cake day u sneakpeekbot!
Emo Labels: 'excitement'
Mapping Col: Joy
--------------------
Text: was he rejected because of his methodology or because of the content of his studies?
Emo Labels: 'confusion'
Mapping Col: Surprise
--------------------
Text: thanks, i agree
Emo Labels: 'approval', 'gratitude'
Mapping Col: Trust
--------------------
Text: why would you doubt it dumbass?
Emo Labels: 'anger', 'disgust'
Mapping Col: Disapproval, Annoyance
--------------------
```

Figure 7. Semantic Mapping

STEP 4: SELF-ATTENTION

After mapping emotions with semantic models, the subsequent step is to introduce the Self-Attention and Multi-Head Attention mechanisms to examine long-range dependencies and extract complex patterns of emotional expression. Attention mechanisms are especially valuable in NLP-based depression analysis because they enable the model to pay attention to significant words or phrases that affect emotions. Self-attention allows the model to assign different weights to words depending on their context relevance. In detecting depression, some words will be highly indicatives distress while others are less important. The attention mechanism assigns higher weights to emotionally charged words like "hopeless," "lonely," or "exhausted," ensuring that they contribute more to emotion classification. This helps detect emotions linked to depression symptoms in text responses.

Example: When a patient shares a statement like: "*I feel lost. No motivation. Nothing excites me. But I try to be happy.*" The Self-Attention mechanism assigns higher weights (Figure 7) to words strongly associated with negative emotions (e.g., *lost, no motivation, nothing*), which are critical for depression detection.

Figure 8. Self-Attention Weights

Words like *"lost"*, *"no motivation"*, and "nothing" receive high attention scores, meaning the model focuses on these terms when classifying the emotion. Words like "happy" have a low attention score because they contradict the dominant negative sentiment is shown in Table 3.

Table 3: Interpretation of Self-Attention Mechanism

Word	Attention Score	Interpretation
Lost	0.91	Strong signal for distress
no motivation	0.88	Indicates lack of energy, a depression symptom
Nothing	0.85	Suggests hopelessness
Happy	0.40	Lower attention as it contradicts the overall emotion

STEP 5: MULTI-HEAD ATTENTION FOR COMPREHENSIVE EMOTION ANALYSIS

By combining different attention heads, the model gains a holistic understanding of the patient's emotional state. Multi-Head Attention extends Self-Attention by using multiple attention heads to focus on different parts of the input text simultaneously. Imagine a patient shares the statement:

Example: *"I feel lost. No motivation. Nothing excites me. But I try to be happy."*

Instead of one attention mechanism, Multi-Head Attention uses multiple attention heads, each focusing on a different aspect of the sentence:

Head 1: Emotional Words Focus

Detects strong emotional indicators (e.g., lost, no motivation, nothing). Helps identify depressive tendencies.

Head 2: Context Understanding

Recognizes how words relate to each other. Understands that "I try to be happy" might be a contrast to negativity rather than a positive emotion.

Head 3: Sentiment & Sarcasm Detection

Differentiates genuine sadness from sarcasm. If someone says, "Oh great, I'm totally fine!" → This head detects sarcasm. Figure. 9 represents the attention weights from all eight heads of the Multi-Head Attention mechanism applied to an input sentence containing emotionally significant words. Each attention head captures different aspects of the relationships between words, ensuring a more comprehensive understanding of the context. Interestingly, strongly emotional words like "lost", "no motivation", and "nothing" get higher attention scores across several heads, suggesting that they are significant in identifying depressive language. Certain heads concentrate on self-attention, where words mainly focus on themselves, but others pick up on the word dependencies to ensure the model identifies relationships between the emotional triggers.

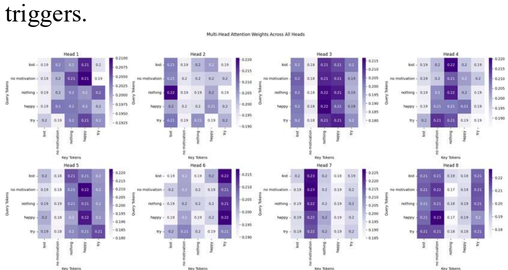

Figure 9: Multi-Head Attention Weights

In Figure 9, Heads 7 and 8 specifically stress depressive words and give more weight to terms such as "no motivation" and "lost", which are significant indicators in detecting depression. Head 8 also emphasizes "happy", which may feel contrast in emotions and would be beneficial for distinguishing true from covered emotional states. The diversified patterns of attention for every head ensure that both distinctive significance of individual words and word-word relationships are captured well. This visualization supports that multi-head attention enhances emotion recognition by identifying crucial emotional stimuli, semantic relations, and capturing contrast in context. This ability is essential to help improve the accuracy of AI-based depression detection models as it helps achieve a more nuanced representation of emotional cues in text. Real-time social media emotion recognition is a significant component of mental health assessments, providing medical professionals with an AI-supported approach to the identification of high-risk individuals.

5. CONCLUSION AND FUTURE DIRECTIONS

Real-time emotion detection on social media is critical in mental health assessment since it provides clinicians with an AI-driven means of identifying at-risk individuals. The Emotion-Aware AI for Mental Health (EAMH) system makes it possible to track emotions and precursors to distress continuously through the analysis of posts made by users. Using datasets like GoEmotions and advanced self-attention and multi-head attention, AI systems can process emotional trends in real time and react appropriately. Preprocessing keeps data representative by removing duplicates and redundant metadata, and BERT tokenization preserves contextual connections. Context comprehension enhances emotion analysis further by recognizing sentiment reversals, sarcasm, and strength of emotions, reducing misinterpretation. Semantic mapping puts recognized emotions into psychological frameworks to maximize the clinical utility of mental health assessments. With attention-based AI included, EAMH provides a structured, interpretable model for emotion recognition with relevant information for clinicians to act upon before mental distress results in negative outcomes. Self-attention and multi-head attention provide the model with the capability to assign varying weights of importance to emotionally charged words, allowing it to track depressive patterns of language effectively. Recording vital emotional cues and context-specific connotations allows

By continuously monitoring mental health, AI can identify emotions to identify depression, anxiety, or suicidal ideation at an early stage, making social media an important medium for maintaining your mental health. In addition, the use of Explainable AI (XAI) also ensures that there is transparency, allowing clinicians to understand model predictions and improve personalized treatment plans. Although the EAMH model displays high capability in identifying depression-related emotions, it possesses some limitations. Furthermore, mapping GoEmotions labels to psychological frameworks such as Ekman's and Plutchik's theories may lead to semantic mismatches, potentially affecting the precision of emotion categorization. Additionally, the current approach is limited to text-based analysis, excluding multimodal emotional signals such as tone, facial expressions, or physiological cues, which are crucial for accurate affective computing. Future extensions of this work will focus on incorporating multimodal data sources (e.g., audio, video, and biometric signals), training on clinically annotated datasets, and refining the integration of psychological emotion models to enhance the model's robustness, diagnostic relevance, and applicability in real-world mental health assessments.

Acknowledgements. I (SUDHANDRADEVI P), express my gratitude for providing financial support under the fellowship of the University Research Fellowship, Bharathiar University, Coimbatore.
Disclosure of Interests No competing interests.

REFERRENCE

Akbar Karimi, A., Aghajanzadeh, S., Joorabloo, N., Morency, L.-P., & Wong, D. F. (2020). Adapting BERT for ABSA: A study on parallel and

hierarchical aggregation methods. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP 2020)*, 4630–4640. https://doi.org/10.18653/v1/2020.emnlp-main.375
[

Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. *3rd International Conference on Learning Representations (ICLR 2015)*. https://doi.org/10.48550/arXiv.1409.0473
[[

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., & Dhariwal, P. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems (NeurIPS 2020)*, 33, 1877–1901. https://arxiv.org/abs/2005.14165

Calvo, R. A., D'Mello, S., Gratch, J., & Kappas, A. (2017). *The Oxford handbook of affective computing*. Oxford University Press.

Cambria, E., Poria, S., Gelbukh, A., & Thelwall, M. (2017). Sentiment analysis is a big suitcase. *IEEE Intelligent Systems, 32*(6), 74–80. https://doi.org/10.1109/MIS.2017.4531228

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2019)*, 4171–4186. https://doi.org/10.18653/v1/N19-1423

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., … Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *International Conference on Learning Representations (ICLR 2020)*. https://arxiv.org/abs/2010.11929

Dutta, S., & Ganapathy, S. (2023). A hierarchical cross-attention model for multi-modal sentiment analysis. *IEEE Transactions on Affective Computing*. https://doi.org/10.1109/TAFFC.2023.3284561

Fan, X., Zhou, X., Liu, Y., & Sun, S. (2023). Multi-granularity attention network for emotion recognition in conversation. *Neural Networks, 165*, 76–88. https://doi.org/10.1016/j.neunet.2023.04.015

Felbo, B., Mislove, A., Søgaard, A., Rahwan, I., & Lehmann, S. (2017). Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion, and sarcasm. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP 2017)*, 1615–1625. https://doi.org/10.18653/v1/D17-1169

Hu Xu, X., Liu, C., Bing, L., & Liu, B. (2020). Aspect-based sentiment classification with BERT. *Proceedings of the 2020 AAAI Conference on Artificial Intelligence (AAAI 2020)*, 34(05), 9195–9202. https://doi.org/10.1609/aaai.v34i05.6399

Kim, Y. (2014). Convolutional neural networks for sentence classification. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1746–1751. doi.org/10.3115/v1/D14-1181

Mickel Hoang, M., & Motlicek, P. (2019). Aspect-based sentiment analysis using BERT. *Proceedings of the 2019 Conference of the Association for Computational Linguistics (ACL 2019)*, 6280–6286. https://doi.org/10.18653/v1/P19-1629

Mingke Xu, X., Lin, Y., & Li, Z. (2021). Head fusion method based on multi-head attention for speech emotion recognition. *IEEE Transactions on Affective Computing, 12*(3), 482–493. https://doi.org/10.1109/TAFFC.2021.3079356

Nediyanchath, N.A., S.G. Koolagudi, and K.S. Rao. 2020. Multi-head self-attention with position embedding for speech emotion recognition. IEEE Access 8: 182351–182362. https://doi.org/10.1109/ACCESS.2020.3028194

Pecoraro, G., Zanlorensi, L. A., Gallo, I., & De Gregorio, M. (2021). Local multi-head channel self-attention for facial expression recognition. *IEEE Transactions on Image Processing, 30*, 8448–8461. https://doi.org/10.1109/TIP.2021.3118391

Ping Ji, J., & Xiaohong Liu, X. (2020). A transformer-based approach to speech emotion recognition with multi-head attention and position weighting. *IEEE Access, 8*, 82845–82856. https://doi.org/10.1109/ACCESS.2020.2991582

Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2020). A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion, 37*, 98–125. https://doi.org/10.1016/j.inffus.2017.02.003

Schuller, B., Batliner, A., Steidl, S., & Vogt, T. (2021). Speech emotion recognition: Two decades in a nutshell, benchmarks, and ongoing trends. *Communications of the ACM, 64*(8), 62–71. https://doi.org/10.1145/3461641

Shen, J., Rudzicz, F., Begum, M., & Yancheva, M. (2020). Detecting anxiety through self-reported speech patterns and language use. *Frontiers in Psychology, 11*, 2311. https://doi.org/10.3389/fpsyg.2020.564753

Shen, W., & Xu, T. (2023). Multi-headed channel attention for facial expression recognition with ResNet18.[*Pattern Recognition Letters, 163*, 78–86. https://doi.org/10.1016/j.patrec.2023.02.008

Siriwardhana, S., Kaluarachchi, T., & Fernando, A. (2020). A self-supervised transformer-based fusion model for multi-modal emotion recognition. *Proceedings of the 2020 International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2020)*, 8064–8068. https://doi.org/10.1109/ICASSP40776.2020.9054429

Song, J. (2023). Vision Transformers for facial expression recognition: A study on local and global pattern recognition. *IEEE Transactions on Affective Computing, 14*(2), 145–157. https://doi.org/10.1109/TAFFC.2023.3245123

Tzirakis, P., Trigeorgis, G., Nicolaou, M. A., Schuller, B., & Zafeiriou, S. (2017). End-to-end multimodal emotion recognition using deep neural networks. *IEEE Journal of Selected Topics in Signal Processing, 11*(8), 1301–1309. https://doi.org/10.1109/JSTSP.2017.2764438

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems (NeurIPS)*, 30, 5998–6008. https://doi.org/10.48550/arXiv.1706.03762

Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2017). Graph attention networks. *Proceedings of the International Conference on Learning Representations (ICLR 2018)*. https://arxiv.org/abs/1710.10903

Wen, Z., Han, J., & Zhang, W. (2021). A two-stream global-guided attention network for facial expression recognition. *Pattern Recognition, 119*, 108067. https://doi.org/10.1016/j.patcog.2021.108067

Wu, H., Zhang, S., & Chen, Y. (2023). Progressive tri-modal attention network for multi-modal emotion recognition. *Neural Computing and Applications, 35*, 12841–12853.https://doi.org/10.1007/s00521-023-07892-5

Xiao, L., Chen, H., & Xu, Y. (2024). Improving VGG16 with self-guided attention and feature fusion for facial expression recognition. *Pattern Recognition Letters, 169*, 45–57. https://doi.org/10.1016/j.patrec.2024.01.007

Xinxin Jiao, J., & Zhang, T. (2024). Multi-spatial fusion and hierarchical cooperative attention for speech emotion recognition. *IEEE Transactions on Affective Computing*. https://doi.org/10.1109/TAFFC.2024.3284711

Yu, C., & Bai, Z. (2021). Visual self-attention network for facial expression recognition under environmental disturbances. *IEEE Transactions on Image Processing, 30*, 8934–8945. https://doi.org/10.1109/TIP.2021.3125572

Zhang, Y., Wang, H., & Li, M. (2023). DualGATs: Speaker-aware discourse structure modeling with dual graph attention networks. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023)*, 2950–2961. https://doi.org/10.18653/v1/2023.emnlp-main.400

Zhou, C., Sun, C., Liu, Z., & Lau, F. C. M. (2016). A C-LSTM neural network for text classification. *arXiv preprint arXiv:1511.08630*. https://doi.org/10.48550/arXiv.1511.08630

Zhou, Q., Liu, F., & Tang, J. (2022). Regional self-attention convolutional neural network for facial expression recognition. *IEEE Transactions on Affective Computing, 13*(3), 492–505. https://doi.org/10.1109/TAFFC.2022.3141283

Books

Goleman, D. (1995). *Emotional intelligence: Why it can matter more than IQ*. Bantam Books.