Lightweight Multi-scale Convolutional Framework with Context-augmented Embeddings for Efficient Aspect-based Sentiment Analysis

Susmitha Mekala [1,2,*] and Shaik Razia [1]

[1] Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation (Deemed to be University), Andhra Pradesh, India
[2] Department of Information Technology, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India
Email: 183030081@kluniversity.in (S.M.); skrazia@kluniversity.in (S.R.)
*Corresponding author

***Abstract*—Aspect-Based Sentiment Analysis (ABSA) is an opinion mining method that finds fine-grained opinions, which are aspect terms and their associated sentiment polarities in user-generated text. Current methods are characterized by high computational cost, lack of sensitivity to subwords and error propagation between sequential subtasks. It is suggested in this paper that a new lightweight and end-to-end ABSA architecture, which includes three closely linked components (i) a Context-Augmented Subword Embedding (ii) mechanism to augment static embeddings with contextual bias, polarity modifier cues, and domain-relevant reinforcement, (iii) a Hierarchical Multi-Scale Convolutional Neural Network (HMSCNN) to learn a wider range of local, intermediate, and long-range sentiment dependencies, using multi-dilated parallel convolutions. The extensive experiments with the SemEval-2014 Restaurant and Laptop benchmark datasets have shown that the proposed framework will run to achieve a Span-level Aspect Term Extraction (ATE) F1-score of 83.6%, Attention-based Sentiment Classification (ASC) accuracy of 86.4% with Macro-F1 of 84.1% and Joint F1-score of 76.2%, outperforming Convolutional Neural Networks-Recurrent Neural Networks (CNN-RNN) and recurrent baselines by 4.4%–13.3%. The additive contribution of each component is confirmed with aid of ablation studies. These findings indicate that the suggested framework can provide a good compromise between representational richness and computational efficiency and can be used in real-time and resource-constrained fine-grained sentiment analysis tasks.**

***Keywords*—aspect-based sentiment analysis, aspect term extraction, multi-scale convolutional neural network, context-augmented embeddings, aspect-guided attention, Semantic Evaluation (SemEval-2014), lightweight natural language processing**

I. Introduction

The proliferation of web technologies and online shopping websites has created a new kind of user-generated textual content, online reviews, feedback posts, social media interactions, etc. These texts contain useful emotion-carrying information that represents user opinions, preferences, and experiences making them the most resourceful source of information for businesses, policymakers, and researchers [1]. Despite being handy with document/sentence polarities, conventional sentiment analysis methods only offer high-level views of opinions [2]. Practically, one review can simultaneously be positive in relation to some aspects and negative in relation to others, requiring more subtle analysis.

Aspect-Based Sentiment Analysis (ABSA) has emerged as a niche area of Natural Language Processing (NLP) to support this need, by identifying sentiment polarity in respect to aspects or features of an entity [3, 4]. Aspect-Based Sentiment Analysis (ABSA) divides opinions into aspect-sentiment pairs and thus allows the user feedback to be interpreted with a greater sense of nuance and facilitate fine-grained downstream analytics. Benchmark initiatives have also greatly spurred the development of ABSA, with the semeval-2014 Task 4 shared evaluation, providing manually annotated datasets of restaurant and laptop reviews as a standardized evaluation of aspect term extraction and aspect-level sentiment classification [5].

Since the inception of these benchmarks, ABSA approaches have evolved through rule-based practices, classical machine learning architectures, deep neural networks, and, finally, transformer-based architectures [6]. Initial techniques relied on handcrafted syntactic and semantic properties, limiting scalability and cross-domain adaptability. Deep learning techniques provided automatic feature learning and better generalization. Transformer models such as Bidirectional Encoder Representations from Transformers (BERT) and its variants have improved performance by incorporating complex contextual dependencies and implicit sentiment

Manuscript received March 23, 2026; revised June 3, 2026; accepted June 25, 2026; published November 10, 2026.

cues. However, these performance benefits have often been achieved at the expense of significantly increased computational complexity, memory footprint, and inference latency, posing practical constraints for real-time or resource-constrained deployment.

Even with tremendous improvements, several core issues remain unaddressed in ABSA research. First, most current models find it challenging to model fine-grained sentiment indicators at the subword level, especially when there are implicit cues, sentiment modifiers, or subtle polarity expressions [7]. Second, the computational overhead of transformer-based architecture limits their applicability in deployment-oriented scenarios. Third, traditional ABSA pipeline systems that define extraction aspect and sentiment classification as separate sequential tasks create cascading error effects and do not allow joint optimization.

To counter these flaws, this paper proposes a new lightweight, multicomponent ABSA framework that achieves high accuracy while remaining computationally efficient. The proposed approach builds upon enriched embedding techniques to capture local and domain-sensitive sentiment indicators, multi-scale contextual representation learning to support fine-grained interpretation, and adaptive task-specific mechanisms for both extraction and classification. Extensive tests at the SemEval-2014 Restaurant and Laptop datasets prove that the present framework increases the accuracy of aspect boundary detection, reduces the proliferation of subtask error, and has superior performance to transformer-intensive models, while maintaining efficiency and scalability. Related approaches leveraging deep learning for sentiment analysis have demonstrated the potential of embedding-based and CNN-based methods while local context focus mechanisms and context-aware dynamic embeddings have established strong foundations for efficient aspect-level inference.

The main contributions of this work are summarized as follows:

- A Context-Augmented Subword Embedding (CASE) mechanism is introduced that enriches static fastText vectors through three additive signals contextual bias, polarity modifier encoding, and domain reinforcement within a single lightweight transformation requiring no additional encoder pass.
- A Hierarchical Multi-Scale Convolutional Neural Network (HMSCNN) is proposed that captures sentiment-relevant patterns at local, intermediate, and long-range contextual scales through three parallel dilated convolutional branches. The architecture is fully convolutional, recurrence-free, and inherently parallelizable.
- A Context-Adaptive Aspect Term Extraction (ATE) head is developed that applies sigmoid-gated token relevance scoring before Beginning, Inside, Outside (BIO) sequence labelling, dynamically suppressing non-aspect tokens and improving aspect boundary detection, particularly for multi-word expressions.
- An Aspect-Guided Attention-based Sentiment Classification (ASC) head is designed in which attention weights are explicitly conditioned on the extracted aspect representation, preventing surrounding polarized context from distorting the sentiment prediction for the target aspect.
- All four components are optimized jointly through a single combined training objective, enabling multi-task gradient flow through the shared HMSCNN backbone and ensuring representations are simultaneously informative for both aspect extraction and sentiment classification.
- Comprehensive experiments on the SemEval-2014 Restaurant and Laptop datasets covering span-level Aspect Term Extraction (ATE), token-level Beginning, Inside, Outside (BIO) tagging, boundary analysis, aspect-length stratification, domain-wise robustness, per-class Attention-based Sentiment Classification (ASC), end-to-end joint evaluation, ablation, and convergence analysis demonstrate consistent improvements over five representative baseline models with statistical significance confirmed across multiple independent runs.

The remainder of the paper is organized as follows. Section II reviews related work in ABSA. Section III describes the proposed methodology. Section IV details the experimental setup. Section V presents and analyzes experimental results. Section VI discusses key findings and limitations. Section VII concludes the paper and outlines future research directions.

II. Literature Review

The Aspect-Based Sentiment Analysis (ABSA) has become a core task in the opinion mining of user-generated text at the fine-grained level. In contrast to document- or sentence-level sentiment analysis, ABSA is geared towards detecting sentiment polarity with reference to features of an entity, which allows more accurate and actionable information. Recent literature indicates that there has been sustained progress in contextual representation learning, joint modeling, syntactic reasoning and efficiency-conscious architectures. This section synthesizes the available literature and establishes unresolved gaps that inspire the proposed framework.

A . Task Formulation and Persistent Challenges

In current literature, ABSA is consistently defined as a compound learning problem comprising two interconnected subtasks: Aspect Term Extraction (ATE) and Aspect Sentiment Classification (ASC) [8, 9]. Although SemEval-2014 benchmark datasets remain central to experimental evaluation, multiple studies have found that benchmark-based optimization underestimates the sophistication of real-world sentiment expression. Continuing challenges include the reliable identification of implicit aspects, proper treatment of sentiment modifiers and negation, sensitivity to domain variation, and performance degradation in sentences with multiple, overlapping, or sparsely expressed aspects.

A systematic review of ABSA in Artificial Intelligence Review noted that most recent ABSA systems have achieved performance gains primarily through

architectural scaling rather than deeper linguistic modeling, raising concerns about scalability, interpretability, and deployment feasibility [8]. Bibliometric studies confirm that domain dependence and dataset bias remain serious limitations even for high-performing models [9], so that better model robustness and representational quality are open research issues.

B. Contextual Representation Learning and Practical Constraints

Improvements in ABSA performance have been mostly led by contextual representation learning, and transformer-based encoders have shown excellent ability in capturing long-range dependencies and implicit sentiment information. Fine-tuned contextual encoders have consistently enhanced accuracy in detection and sentiment measurement.

Empirical findings have, however, consistently pointed out that these performance improvements come at significantly larger computational cost, memory cost, and inference latency, making them less applicable in real-time, edge, or resource-constrained deployment contexts [10]. The computational contrast between deep learning frameworks has been studied to better understand deployment trade-offs.

Beyond computational concerns, interpretability has proven to be a major constraint. Under explainable ABSA, studies report that attention distributions acquired by contextual encoders do not always match human reasoning, especially in linguistically complex or multi-aspect sentences [11]. Moreover, contextual models tend to exhibit reduced generalization between domain transfers, motivating the desire for representations that are simultaneously context-sensitive and domain-resilient.

C. Syntax-aware and Graph-based Modeling

In order to improve aspect-opinion correspondence, a number of studies have included syntactic dependency data using graph-based neural networks. These strategies are also the explicit way of modelling interactions between sentiment-carrying expressions and aspect terms by modeling dependency interactions as structured graphs. Dependency-weighted graph attention mechanisms have been effective in enhancing polarity prediction by prioritizing syntactically relevant contextual cues [12]. Joint graph-based models that combine ATE and ASC into single pipelines have also demonstrated improved cross-task consistency and fewer error propagations [13]. Even when effective, graph-based techniques incur extra computational burden and require fine-quality dependency parsing, which is poorly scalable and robust in real-world applications.

D. Efficiency-oriented Architectures and Multi-scale Feature Learning

Following the principle of deployment efficiency, Convolutional Neural Networks (CNNs) have been revisited as practical alternatives for ABSA. CNNs provide inherent parallelism and reduced inference cost while maintaining competitive ability to capture local sentiment patterns. Multi-scale convolutional feature extraction can model both shortrange sentiment cues and broader contextual semantics [14]. The combination of CNN with BiGRU for short-text ABSA has demonstrated strong performance in the applied domain. Despite these benefits, CNN-based ABSA models commonly do not include hierarchical modeling techniques capable of capturing long-range dependencies and adopt fixed convolutional features without adaptive aspect boundary refinement and aspect-directed sentiment aggregation. Context-aware dynamic word embeddings have also been shown to improve aspect term extraction performance, while combinations of CNN and gated recurrent units offer complementary advantages.

E. Robustness and Linguistic Complexity

Recent studies have paid more attention to the nature of cognitive failures of the ABSA models with respect to linguistic constructions. The research proposing the notion of sentiment difficulty states that syntactic ambiguity, implicit sentiment, sarcasm, and overlap are disproportionately affecting the model accuracy loss [15]. Massive review and microblog data analyses also indicate the abundance of informal and domain-specific words, as well as noisy statements that are harmful to existing systems [16]. The results of these research work show that the linguistic complexity increases resilience rivals' high average performance in the curated benchmarks, and models without fine-grained representational flexibility and adaptive cognition biases are exceptionally susceptible in unstructured deployment tests [17].

F. Research Gaps and Motivation

Altogether, the available literature demonstrates a set of significant unresolved research gaps that directly motivate the current work. First, contextual and graph-based models are more effective at aligning aspects and opinions, but usually require more computational resources, making deployment in realistic settings impractical. Second, joint learning models avoid pipeline-based error but are typically based on heavyweight encoders and non-dynamic task interaction mechanisms. Third, current methods largely neglect sentiment enrichment at the subword level, yet this has been proven essential for capturing polarity modifiers, negation, and implicit sentiment. Local context focusses mechanisms provide partial solutions, but no existing lightweight framework comprehensively addresses contextual bias, polarity modification, and domain reinforcement within a single embedding module. These gaps directly motivate the proposed framework.

III. METHODOLOGY

This section introduces the methodology for a fine-grained Aspect-Based Sentiment Analysis (ABSA) framework designed to capture subtle sentiment cues, identify specific sentiment polarities, and maintain lightweight, computationally efficient architecture. The framework addresses the shortcomings noted in current

ABSA systems, namely the insensitivity to sentiment at the subword level, inter-subtask error propagation, and deployment inefficiency. The proposed methodology incorporates four major components: (i) Context-Augmented Subword Embeddings (CASE) for enriched token representation; (ii) a Hierarchical Multi-Scale Convolutional Neural Network (HMSCNN) for contextual feature extraction; (iii) a Context-Adaptive Aspect Term Extraction (ATE) head for precise aspect boundary detection; and (iv) an Aspect-Guided Attention-based Aspect Sentiment Classification (ASC) head for discriminative sentiment inference. The overall system architecture is shown in Fig. 1.

A . System Overview

The pipeline is in a modular form as demonstrated in Fig. 1, with every part of the pipeline having a specific use. Normalization and tokenization Raw input reviews are mapped to enriched representation vectors via the Context-Augmented Subword Embeddings (CASE) mechanism. These embeddings are passed to the HMSCNN backbone. These embeddings are given on the HMSCNN backbone to produce multi-scale token features that are contextual. The resultant shared representations are invoked by two task-specific heads: (i) Context-Adaptive-ATE, this head labels BIO-tag sequences by identifying aspect term spans, and (ii) an Aspect-Guided Attention-based ASC head that predicts sentiment polarity for each identified aspect span. The overall result is structured aspect-sentiment pairs for each input review.

1) Distinctive elements of the proposed framework

The framework is characterized by the close integration of four complementary elements to enhance fine-grained ABSA and computational efficiency together. At the representation level, the CASE module (Fig. 2) augments static embeddings with subword polarity cues, contextual bias, and domain-specific semantic indicators. The HMSCNN backbone (Fig. 3) takes these enriched embeddings and encodes local, intermediate, and long-range contextual patterns through parallel convolutional branches at different dilation rates [18, 19]. The Context-Adaptive ATE module uses dynamic token weighting and multi-scale feature alignment to optimize aspect boundary identification in diverse contextual conditions. The Aspect-Guided Attention-based ASC module selectively emphasizes sentiment-relevant contextual regions conditioned on extracted aspect spans, enabling accurate aspect-specific polarity prediction with reduced inter-subtask error propagation.

The combination of these four elements addresses a specific gap in the literature that existing lightweight ABSA systems leave open: the lack of a framework that jointly handles subword-level sentiment enrichment, multi-scale parallel feature extraction, and aspect-conditioned sentiment aggregation, without relying on recurrent encoders or pre-trained transformer weights. Prior CNN-based approaches address multi-scale extraction but not dynamic aspect gating. Graph-based systems achieve strong aspect-opinion alignment but at substantially higher computational cost. CASE-style embedding augmentation has not been combined with hierarchically dilated CNNs in any previously published ABSA architecture.

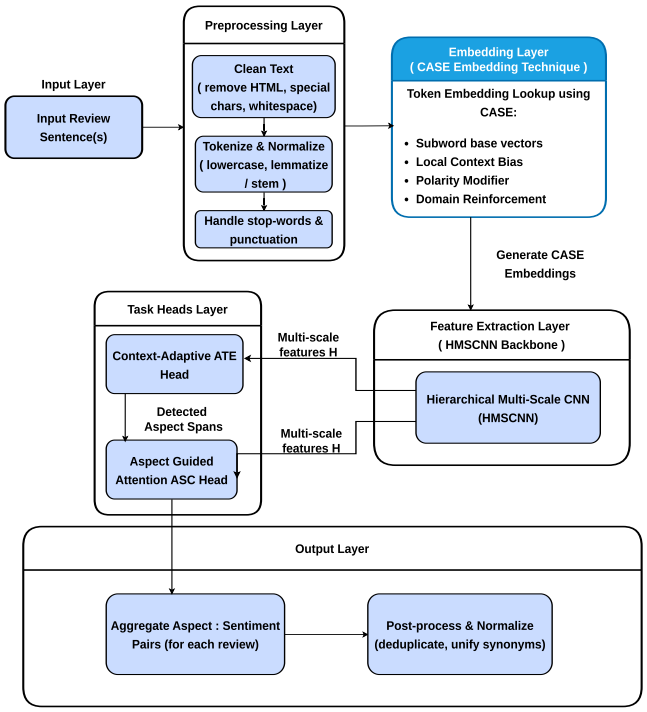


Fig. 1. Layered system architecture of the proposed ABSA framework. The pipeline flows from raw text through CASE enrichment, HMSCNN feature extraction, and dual task-specific heads for ATE (BIO sequence labeling) and ASC (aspect-guided polarity classification).

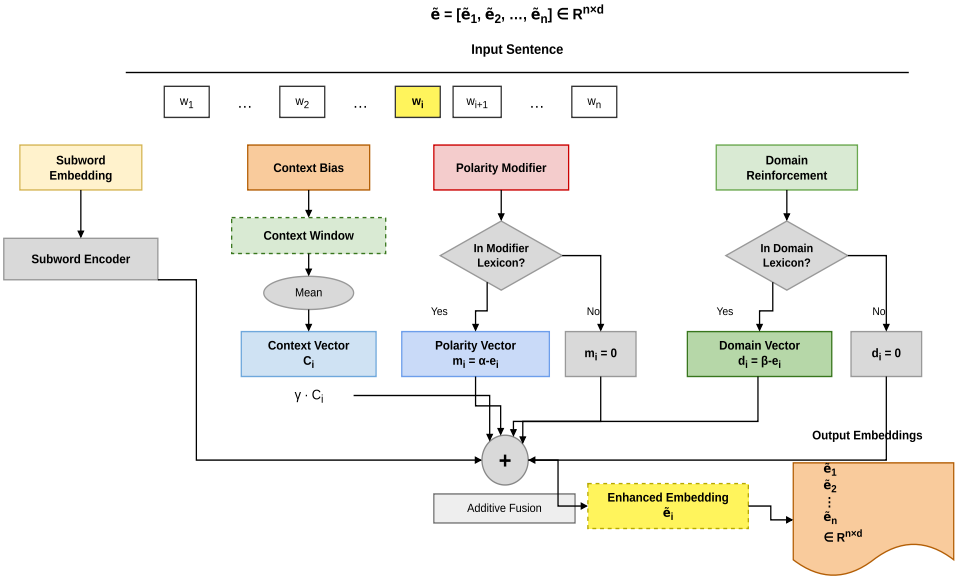


Fig. 2. Conceptual architecture of the Context-Augmented Subword Embedding (CASE) layer. The CASE module fuses three additive components—contextual bias (c_i), polarity modifier (m_i), and domain reinforcement (d_i)—with the base fastText embedding (e_i) to produce an enriched representation $\tilde{e}_i$.

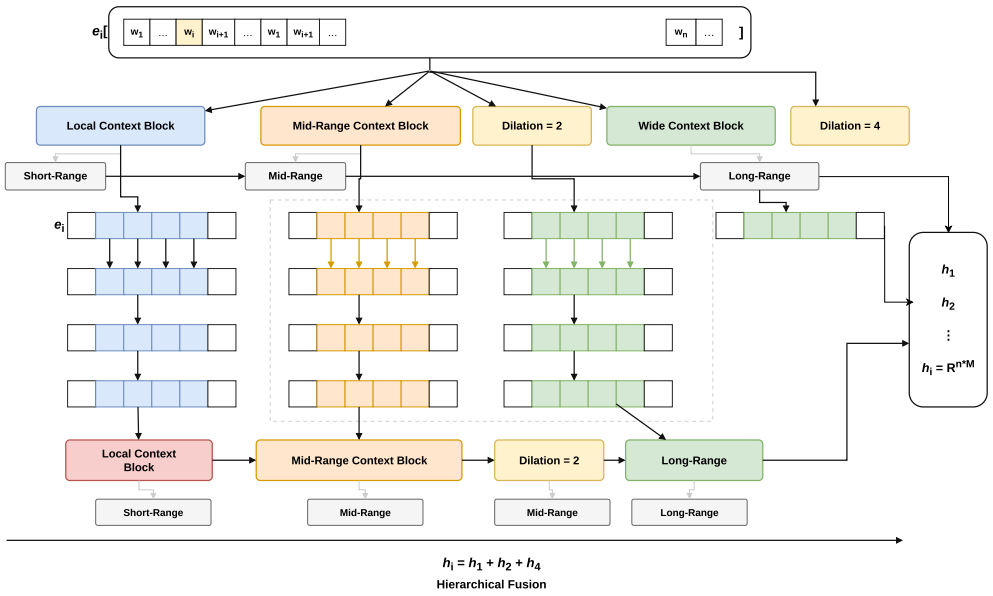


Fig. 3. HMSCNN architecture with hierarchical multi-scale feature extraction. Three parallel dilated convolutional branches Local Context Block (LCB: $d = 1$); Mid-Range Context Block (MCB: $d = 2$); Wide Context Block (WCB: $d = 4$) process the CASE embeddings, and their outputs are concatenated to form the multi-scale feature matrix $H \in \mathbb{R}^{n \times M}$.

B. Input Preprocessing and Token Representation

A raw review document is represented as a list of characters $R = \{c_1, c_2, \ldots, c_m\}$, where c_i denotes the i-th character. A standardized preprocessing function $P(\cdot)$ normalizes the input and removes noise while maintaining sentiment-related information, including lowercasing, normalization of repeated symbols, and removal of non-informative characters. The normalized text is tokenized into a word-level token sequence $T = \{w_1, w_2, \ldots, w_n\}$. Stop words are deliberately preserved because they perform vital functions in sentiment modification through negation, intensification, and contrast constructions.

Each token $w_i \in T$ is initially mapped to a static embedding vector $\mathbf{e}_i \in \mathbb{R}^d$. Subwordlevel sentiment information is incorporated through the CASE function $\mathcal{E}_{\text{CASE}}(\cdot)$, yielding the enriched token representation $x_i = \mathcal{E}_{\text{CASE}}(w_i, \mathcal{C}_i)$, where $\mathcal{C}_i$ denotes the local contextual window surrounding token w_i. The complete token sequence is transformed into an enriched representation matrix $X = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^{n \times d}$, serving as input to the HMSCNN backbone. This design balances representational expressiveness and computational efficiency while preserving sensitivity to nuanced sentiment expressions.

C. Context-Augmented Subword Embedding (CASE)

1) Motivation and design rationale

Static word embeddings such as GloVe [20] and fastText [21] are widely adopted in lightweight ABSA systems because they are computationally efficient and provide broad lexical coverage. However, these embeddings are fixed at training time and therefore insensitive to local contextual variation, sentiment-altering modifiers such as negation and intensification, and domain-specific semantic emphasis all of which are important to fine-grained aspect-level sentiment identification. Fully contextualized embeddings deal with at least some of these drawbacks, at a huge computational cost. The CASE mechanism, as shown in Fig. 2, bridges this gap by augmenting static embeddings with light-weight contextual, sentiment-sensitive, and domain-sensitive information and in addition to ensuring that the efficiency and interpretability of static representations have been preserved, certain specific augmentations that are relevant to sentiment modelling at the aspect level are also introduced.

Lexicon Construction and Domain Adaptation. The polarity modifier lexicon $\mathcal{M}$ used in Eq. (3) is derived from the Multi-Perspective Question Answering (MPQA) opinion lexicon, supplemented with a manually verified list of negation tokens ({not, never, no, neither, nor}) and common intensifiers (very, extremely, highly, incredibly}). The domain lexicons D for Restaurant and Laptop domains are initialized from the SemEval-2014 training annotations, all noun tokens appearing as annotated aspect terms are collected and manually filtered for domain relevance, resulting in 214 tokens for Restaurant and 187 tokens for Laptop. This process requires approximately 2 h of domain-expert review per domain. For adaptation to a new domain, the same pipeline can be applied to any aspect-annotated corpus; where annotations are unavailable, frequency-ranked nouns from domain-specific unlabeled text, combined with cosine-similarity filtering against seed terms, provide a practical alternative. Investigating fully automatic lexicon induction, for example through dependency-parsed aspect term co-occurrence graphs, is left to future work.

2) Formal definition of CASE embeddings

Let an input sentence be represented as a token sequence $W = \{w^1, w^2, \ldots, w_n\}$. Each token w_i is mapped to a base embedding vector $e_i \in \mathbb{R}^d$ via a subword-aware embedding model (fastText). The CASE module enriches e_i through additive fusion of three complementary components:

$$\tilde{e}_i = e_i + \gamma\, c_i + m_i + d_i \quad (1)$$

where $\tilde{e}_i \in \mathbb{R}^d$ is the enriched embedding of token w_i; $e_i \in \mathbb{R}^d$ is its base fast Text embedding; $\gamma \in \mathbb{R}^+$ is a non-negative scalar that controls the weight of the contextual bias component; $c_i \in \mathbb{R}^d$ is the context bias vector encoding local neighborhood information; $m_i \in \mathbb{R}^d$ is the polarity modifier vector, non-zero only when w_i belongs to the modifier lexicon $\mathcal{M}$; and $d_i \in \mathbb{R}^d$ is the domain reinforcement vector, non-zero only when w_i belongs to the domain lexicon D.

3) Contextual bias encoding

To bring in local contextual understanding, a context bias vector c_i is obtained through the summation of owner embeddings around other tokens within a fixed window of size k:

$$c_i = \left(\frac{1}{2k}\right) \Sigma\, e_j,\ \ j \in \{i-k, \ldots, i+k\}, j \neq i \quad (2)$$

where $c_i \in \mathbb{R}^d$ is the context bias vector for token i; k is the window half-size (set to $k = 2$); e_j is the base embedding of neighbour token j; the index set $\{i-k, \ldots, i+k\} \setminus \{i\}$ excludes the token itself; the factor $\frac{1}{2k}$ normalises the sum so that the magnitude of c_i does not grow with k; and γ (Eq. (1)) down-weights c_i's contribution relative to the core embedding e_i.

This formulation wants to capture short-range syntactic and semantic cues that make a difference in sentiment interpretation. e.g.. contrastive conjunctions or sentiment bearing neighbors. Regular contraction of contextual information is avoided by the scaling factor γ that lowers the contribution of contextual information.

4) Polarity modifier awareness

Sentiment polarity is often altered by linguistic modifiers such as negations and intensifiers. A polarity modifier vector is explicitly introduced to encode such effects:

$$m_i = \begin{cases} \alpha e_i & \text{if } w_i \in \mathcal{M} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

where $m_i \in \mathbb{R}^d$ is the polarity modifier vector; $\mathcal{M}$ is a curated lexicon of sentiment-altering tokens (negations such as "not", "never"; intensifiers such as "very", "extremely"); $\alpha \in (0,1]$ is a scalar modifier strength (set to $\alpha = 0.75$ after grid search over [0.5, 1.0]); and $m_i = 0$ (the zero vector) when $w_i \notin \mathcal{M}$, leaving the base embedding unchanged.

5) Domain reinforcement encoding

Aspect-level sentiment interpretation is extensively domain-sensitive. A domain reinforcement vector is defined to introduce domain-specific emphasis:

$$d_i = \begin{cases} \beta e_i & \text{if } w_i \in \mathcal{D} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where $d_i \in \mathbb{R}^d$ is the domain reinforcement vector; $\mathcal{D}$ is a domain-specific lexicon (e.g., {battery, screen, keyboard} for Laptop; {food, service, ambience} for Restaurant); $\beta \in (0,1]$ is the reinforcement scale (set to $\beta = 0.70$ after grid search); and $d_i = 0$ when $w_i \notin \mathcal{D}$.

6) Output representation and integration

Applying the CASE transformation to all tokens yields the enriched embedding matrix: $\tilde{E} = [\tilde{e}^1, \tilde{e}^2, \ldots, \tilde{e}_n] \in R^{n \times d}$ CASE uses an additive fusion methodology as shown in Fig. 2 which has the advantage of keeping dimensionality intact, the interpretability of every

augmentation term and is quite learnable by further convolutional processing. The resultant embeddings are semantically sensitive, mood-categorical and topic-specific and offer higher input features to the HMSCNN backbone.

D. Hierarchical Multi-Scale Convolutional Neural Network (HMSCNN)

1) Design motivation and overview

The HMSCNN is the feature extraction backbone of the proposed architecture, as shown in Fig. 3. Compared to recurrent and transformer-based encoders, HMSCNN is fully convolutional, enabling parallel computation and reduced parameter complexity. Opinionated text contains sentiment cues at multiple levels: short-range lexical modifiers, mid-range aspect-opinion relations, and long-range contrastive structures. HMSCNN processes the CASE-enhanced embedding matrix $\tilde{E} \in \mathbb{R}^{n \times d}$ through three parallel convolutional branches differentiated by dilation factor, followed by hierarchical feature fusion.

2) Multi-scale convolutional blocks

Local Context Block (LCB): Employs dilation rate $d = 1$ to concentrate on adjacent token interactions capturing short-range sentiment cues:

$$h_i^{(1)} = ReLU\left(\sum_{k=-K}^{K} w_k^{(1)} \times \tilde{e}_{i+k}\right) \quad (5)$$

Mid-Range Context Block (MCB): Applies dilation rate $d = 2$ to capture broader syntactic sentiment structures over moderate distances:

$$h_i^{(2)} = ReLU\left(\sum_{k=-K}^{K} w_k^{(2)} \times \tilde{e}_{i+2k}\right) \quad (6)$$

Wide Context Block (WCB): Employs dilation rate $d = 4$ to explain long-range semantic dependencies needed for contrastive or concessive sentiment structures:

$$h_i^{(4)} = ReLU\left(\sum_{k=-K}^{K} w_k^{(4)} \times \tilde{e}_{i+4k}\right) \quad (7)$$

In Eqs. (5)–(7), $h_{\mathrm{i}}^{(1)}, h_{\mathrm{i}}^{(2)}, h_{\mathrm{i}}^{(4)} \in \mathbb{R}^{\{C_{out}\}}$ are the output feature vectors of the Local, Mid-Range, and Wide Context Blocks for token *i*, respectively; $C_{out} = 200$ is the number of output channels per block (so the concatenated feature has M = 600 dimensions); $w_{\mathrm{k}}^{(\mathrm{s})} \in \mathbb{R}^{\{d \times C_{out}\}}$ is the kernel weight matrix for scale $s \in \{1, 2, 4\}$ and position k; $\tilde{\mathrm{e}}_{\mathrm{i+sk}} \in \mathbb{R}^{\mathrm{d}}$ is the CASE-enriched embedding of the token at dilated position $i + sk$; $K = 3$ is the half-width of the kernel (giving a kernel span of $2K + 1 = 7$ positions per branch); $d = 1, 2, 4$ is the dilation rate for LCB, MCB, WCB respectively, controlling the effective receptive field; and ReLU(·) is the rectified linear unit activation applied element-wise.

3) Role of dilation in context modeling

Dilation enables convolutional filters to increase the receptive fields without added computation and without added kernel size. With the proposed design, d = 1 represents the dense and immediate sentiment response; d = 2 represents the intermediate structural response to sentiment; and d = 4 represents the sparse and long-distance sentiment dependence. This multi-dilated approach is made to be such that the model is light weight yet the sensitivity to local sentiment intensity as well as global semantic structure is taken care of.

4) Hierarchical fusion of multi-scale features

The convolutional block will generate a unique feature of every token at a given contextual scale. Those are hierarchically fused by means of concatenation:

$$h_{\mathrm{i}} = h_{\mathrm{i}}^{(1)} \oplus h_{\mathrm{i}}^{(2)} \oplus h_{\mathrm{i}}^{(4)} \quad (8)$$

where $h_i \in \mathbb{R}^{\mathrm{M}}$ is the fused multi-scale representation of token i; $\oplus$ denotes vector concatenation (not addition); $M = C_{out} \times 3 = 600$ is the dimensionality of h_{i} after concatenating the three branch outputs; and $H = [h^1, h^2, \ldots, h_{\mathrm{n}}] \in \mathbb{R}^{\{n \times M\}}$ is the resulting feature matrix passed to the task heads.

E. Context-Adaptive Task Heads

The task head layer uses two sequentially acting modules, (i) a Context-Adaptive Aspect Term Extraction (ATE) head, and (ii) an Aspect-Guided Attention-based Aspect Sentiment Classification (ASC) head. Both modules even rely on the shared multi-scale contextual representations of the HMSCNN backbone to assist in joint but specialized perception of aspect boundary detection and aspect sentiment inference.

1) Context-adaptive Aspect Term Extraction (ATE) head

Context-Adaptive ATE head determines explicit terms boundaries of the aspects by dynamically weighting representations according to context. The context relevance score is computed in each token *i* as:

$$\alpha_{\mathrm{i}} = \sigma(W\mathrm{c}\, h_{\mathrm{i}} + b\mathrm{c}) \quad (9)$$

where $\sigma(\cdot)$ denotes the sigmoid activation and $W\mathrm{c}, b\mathrm{c}$ are learnable parameters. The token representation is then adaptively reweighted:

$$\bar{h}_{\mathrm{i}} = \alpha_{\mathrm{i}} \times h_{\mathrm{i}} \quad (10)$$

Aspect term prediction is formulated as a sequence labeling task using BIO tagging. The reweighted features are passed through a softmax classifier:

$$\hat{\mathrm{Y}}_{\mathrm{i}} = softmax\left(W_{ATE}\bar{h}_{\mathrm{i}} + b_{ATE}\right) \quad (11)$$

Contiguous BIO-labeled tokens are grouped to form single- or multi-word aspect spans. This adaptive formulation improves boundary detection accuracy and reduces sensitivity to contextual variation. Through Sections III, $\sigma(\cdot)$ denotes the element-wise sigmoid function $\sigma(x) = \frac{1}{1+e^{\{-x\}}}$, and softmax $(z)_{\mathrm{k}} = \frac{e^{\{z_{\mathrm{k}}\}}}{\Sigma_{\mathrm{j}}\, e^{\{z_{\mathrm{j}}\}}}$ denotes the standard softmax mapping from $\mathbb{R}^K$ to the K-dimensional probability simplex.

In Eqs. (9)–(11), $\alpha_{\mathrm{i}} \in (0, 1)$ is the context relevance score for token i, computed by a sigmoid gate; $\sigma(\cdot)$ is the sigmoid function, mapping any real value to (0, 1); $Wc \in$

$\mathbb{R}^{\{1\times M\}}$ and $bc \in \mathbb{R}$ are learnable weight vector and bias for the gating layer; $\bar{h}_i \in \mathbb{R}^M$ is the context-weighted token representation after reweighting by α_i; $\hat{y}_i \in \mathbb{R}^{\{|C|\}}$ is the predicted BIO tag probability distribution over tag classes $C = \{B, I, O\}$; $W_{ATE} \in \mathbb{R}^{\{|C|\times M\}}$ and $b_{ATE} \in R^{\{|C|\}}$ are the classification layer parameters; and $softmax(\cdot)$ normalises the logits to a valid probability distribution.

Algorithm 1: Context-Adaptive Aspect Term Extraction (ATE)

Input: $H = [h^1, \ldots, h_n] \in \mathbb{R}^n \times M$ multi-scale feature matrix (output of HMSCNN)
$Wc \in \mathbb{R}^{\{1\times M\}}, bc \in \mathbb{R}$ gating layer parameters
$W_{ATE} \in \mathbb{R}^{\{|C|\times M\}}, b_{ATE} \in \mathbb{R}^{\{|C|\}}$ BIO classifier parameters
Output: A — set of extracted aspect term spans {[s_1,e_1], [s_2,e_2], …}
Step 1 :Context Relevance Scoring (Eq. 9):
for $i = 1$ to n: $\alpha_i \leftarrow \sigma(Wc \cdot h_i + bc)$
Step 2 :Adaptive Token Reweighting (Eq. 10):
for $i = 1\ to\ n$: $\bar{h}_i \leftarrow \alpha_i \cdot h_i$
Step 3 :Sequence Label Prediction (Eq. 11):
for $i = 1\ to\ n$: $\hat{y}_i \leftarrow softmax(W_{ATE} \cdot \bar{h}_i + b_{ATE})$
Step 4: Token Label Assignment:
for $i = 1\ to\ n$: $y_i \leftarrow argmax(\hat{y}_i)\ \{y_i \in \{B, I, O\}\}$
Step 5: Span Construction:
A ← ExtractSpans({y_1, …, y_n}) [merge contiguous B/I tokens into spans]
Return A

2) Aspect-guided attention-based ASC head

The ASC head performs sentiment classification conditioned explicitly on the aspect spans identified by the ATE module. For a detected aspect span spanning tokens *s* to *e*, an aspect representation vector is computed via mean pooling:

$$v_{asp} = \left(\frac{1}{e-s+1}\right)\sum_{i=s}^{e} h_i \tag{12}$$

Aspect-guided attention weights are computed for each sentence token, where W_a is a trainable attention matrix:

$$\beta_i = \frac{\exp\left(h_i{}^T W_a\, v_{asp}\right)}{\Sigma_j \exp\left(h_j{}^T W_a\, v_{asp}\right)} \tag{13}$$

An aspect-aware context vector is then computed:

$$v_{ctx} = \Sigma_i \beta_i \times h_i \tag{14}$$

The final sentiment representation concatenates the aspect and context vectors:

$$z = v_{asp} \oplus v_{ctx} \tag{15}$$

Sentiment polarity is predicted via a softmax classifier:

$$p = softmax(W_{ASC} z + b_{ASC}) \tag{16}$$

In Eqs. (12)–(16), *s* and *e* are the start and end token indices of the detected aspect span; $v_{asp} \in \mathbb{R}^M$ is the aspect representation vector obtained by mean-pooling span token features; $\beta_i \in (0,1)$ is the attention weight assigned to token i, reflecting its relevance to the aspect; $W_a \in \mathbb{R}^{\{M\times M\}}$ is a trainable bilinear attention matrix; $h_i{}^T W_a\, v_{asp}$ computes a scalar compatibility score between token *i* and the aspect vector; $v_{ctx} \in \mathbb{R}^M$ is the aspect-conditioned context vector, a weighted sum of all token features; $z \in \mathbb{R}^{\{2M\}}$ is the final sentiment representation formed by concatenating v_{asp} and v_{ctx}; $W_{ASC} \in \mathbb{R}^{\{|S|\times 2M\}}$ and $b_{ASC} \in \mathbb{R}^{\{|S|\}}$ are parameters of the sentiment classification head; $p \in \mathbb{R}\text{^}\{|S|\}$ is the predicted sentiment probability vector over classes S = {positive, negative, neutral}.

The predicted sentiment label corresponds to the class with highest probability. By combining a context-adaptive ATE head and an aspect-guided attention-based ASC head, the task heads layer enables accurate aspect boundary detection and discriminative sentiment classification in an end-to-end manner.

F. Joint Training and Optimization Strategy

The proposed framework is trained in a joint learning paradigm in which ATE and ASC are optimized simultaneously within a unified architecture. This joint optimization allows the model to learn shared multi-scale contextual representations informative for both subtasks, while retaining task-specific discrimination through dedicated prediction heads. By coupling aspect detection and sentiment inference during training, the framework minimizes the error propagation inherent in pipeline-based ABSA systems and facilitates coherent end-to-end learning.

1) Loss formulation

The ATE loss is defined as the token-level cross-entropy loss over BIO tag classes *C*:

$$L_{ATE} = -\left(\frac{1}{n}\right)\Sigma_i \Sigma_{\{c\in C\}} y_{\{i,c\}} \log \hat{y}_{\{i,c\}} \tag{17}$$

The ASC loss is defined as the cross-entropy loss over sentiment polarity classes *S*:

$$L_{ASC} = -\Sigma_{\{k\in S\}} y_k{}^{sent} \log p_k \tag{18}$$

The overall training objective combines both task-specific losses:

$$L_{total} = \lambda\, L_{ATE} + (1-\lambda) L_{ASC} \tag{19}$$

In Eqs. (17)–(19), L_{ATE} is the token-level cross-entropy loss over BIO tag classes C for all n tokens; $y_i, c \in \{0,1\}$ is the ground-truth indicator for token *i* and BIO class c; $\hat{y}_i, c \in (0,1)$ is the predicted probability for token *i* and class c; L_{ASC} is the cross-entropy loss over sentiment polarity classes *S* for all detected aspects; $y_k^{sent} \in \{0,1\}$ is the ground-truth sentiment label for aspect k; $p_k \in (0,1)$ is the predicted probability of the correct sentiment class for aspect k; $\lambda = 0.55 \in [0,1]$ is the task-weighting hyperparameter, tuned on the validation set; and L_{total} is the unified objective minimised during joint training.

2) Optimization details

Model parameters were optimized using the Adam optimizer, with task-specific learning rates set to 2×10^{-4} for the Aspect Term Extraction (ATE) task and 1.5×10^{-4} for the Aspect Sentiment Classification (ASC) task. A dropout following every convolutional block and L2 weight decay are used to accomplish regularization. Its convolutional backbone and minimal CASE embedding layer make the layer more amenable to a consistently stable training convergence at significantly lower memory overhead rates than transformer-based models. An example is the use of early stopping due to validation loss as being used to avoid overfitting.

IV. Experimental Setup

This section outlines the experimental design used to test the effectiveness and strength of the suggested ABSA framework in terms of the description and division of datasets, model configurations, testing procedures, and baseline models.

A. Datasets and Data Splits

Experiments are conducted on the SemEval-2014 Task 4 benchmark datasets [22]. The task provides domain specific review data manually annotated with aspect term boundaries and sentiment polarity labels across three classes: positive, negative, and neutral. Two domains are examined: restaurant reviews and laptop reviews. Table I summarizes the data set statistics.

TABLE I. Dataset Statistics for SemEval-2014 Task 4

Domain	Split	No. of Sentences	No. of Aspect Annotations
Restaurant	Train	3044	3699
Restaurant	Test	800	1134
Laptop	Train	3048	2358
Laptop	Test	800	654

The datasets are in XML format with predefined training and test splits as specified by the SemEval-2014 challenge organizers. Each domain's official training split is further divided into training (80%) and validation (20%) subsets for model development and hyperparameter tuning. Official SemEval test splits are reserved exclusively for final performance evaluation. This partitioning strategy ensures consistency with prior research and enables direct comparison with published baselines.

B. Model Configuration and Hyperparameters

The hyperparameters are determined by the performance in the validation set, and their values remain constant in the two domains, to be able to have a fair comparison and reproducibility. Table II outline summarizes the entire hyperparameter configuration.

C. Evaluation Metrics

Consistent with the aims of the research, the proposed framework is tested on three stages of complementary level of ABSA performance of the Aspect Term Extraction (ATE), Aspect-Level Sentiment Classification (ASC), and joint end-to-end ABSA performance. Each of the metrics follows the official SemEval evaluation protocol.

TABLE II. Hyperparameter Configuration

Hyperparameter	Value	Notes
Embedding dimensionality (d)	300	fastText (subword-aware)
Context window size (k)	2	Validated range: [1, 3]
Polarity modifier scale (α)	0.75	Tuned in [0.5, 1.0]
Domain reinforcement scale (β)	0.70	Fixed after grid search
Contextual contribution weight (γ)	0.30	Balance context vs. semantics
CNN kernel size (K)	3	All blocks
Dilation factors	1, 2, 4	LCB, MCB, WCB
CNN output channels per block	200	Total M = 600 after concat.
Dropout rate	0.30	After each CNN block
ATE learning rate	2×10^{-4}	Adam optimizer
ASC learning rate	1.5×10^{-4}	Adam optimizer
Batch size	32	Both heads
Training epochs (max)	50	Early stopping (patience = 5)
Joint loss weight (λ)	0.55	Tuned on validation set
Random seed	42	For reproducibility

1) Aspect Term Extraction (ATE)

The proposed framework is evaluated at three complementary levels: Aspect Term Extraction (ATE), Aspect-Level Sentiment Classification (ASC), and joint end-to-end ABSA performance. Each metric follows the official SemEval evaluation protocol.

$$Precision_{ATE} = \frac{TP}{(TP + FP)} \tag{20}$$

$$Recall_{ATE} = \frac{TP}{TP+FN} \tag{21}$$

$$F1_{ATE} = 2 \times Precision_{ATE} \times \frac{Recall_{ATE}}{(Precision_{ATE} + Recall_{ATE})} \tag{22}$$

The F1-score is the primary ATE metric, balancing detection of correct aspects and false detections.

Where True Positives (TP) counts predicted aspect spans whose start and end indices exactly match a ground-truth span annotation; False Positives (FP) counts predicted spans with no exactly matching annotation; False Negatives (FN) counts annotated spans not recovered by the model; and $F1_{ATE}$ is the harmonic mean of $Precision_{ATE}$ and $Recall_{ATE}$,giving equal weight to missed aspects and false detections.

2) Aspect-level Sentiment Classification (ASC)

The model's capacity to assign sentiment polarity to extracted aspect terms is evaluated by ASC. Overall classification accuracy and Macro-F1 are reported. Macro-F1 computes the unweighted mean of per-class F1-scores, controlling for possible class imbalance.

3) End-to-end ABSA evaluation

For end-to-end evaluation, a prediction is correct only if both the aspect boundary and sentiment label exactly match the ground truth. Joint F1-score and Exact-Match

Accuracy are reported [23], with the latter defined as the ratio of input sentences for which all aspect-sentiment pairs are correctly predicted [24].

D. Baseline Models

The proposed framework is compared against Four representative baseline models spanning classical machine learning, recurrent, convolutional, and hybrid architectures:

- TD-LSTM [25]: Target-Dependent LSTM, a canonical recurrent baseline that models left and right contexts of a target using separate LSTM encoders.
- CNN-LSTM Hybrid [26]: Combines CNN-based local feature extraction with LSTM-based sequential context modeling, offering a competitive convolutional recurrent reference.
- CNN-BiRNN Hybrid [27]: Short-text ABSA model based on CNN combined with Bidirectional GRU, representing the state of the art in efficient hybrid architectures.
- CRF Baseline [28]: A Conditional Random Field-based sequence labeler for ATE with SVM-based ASC; historically strong feature-engineered reference.

For contextual reference, published results from two representative transformer-based ABSA models evaluated on the same SemEval-2014 benchmark are included in Section V. These figures are drawn directly from the respective original publications and are not re-implemented, since the objective of this work is not to outperform large pre-trained models but to establish a computationally lightweight and deployable alternative at a substantially reduced parameter cost.

E. Computational Complexity Analysis

The CASE embedding layer introduces no learnable parameters, since α, β, and γ are fixed scalar hyperparameters and the lexicons $\mathcal{M}$ and $\mathcal{D}$ are static. The HMSCNN backbone contains three dilated convolutional branches, each with kernel size 7, input dimension $d = 300$, and 200 output channels, giving 420,000 parameters per branch and 1,260,000 across all three. The ATE head contributes approximately 2400 parameters and the ASC bilinear attention matrix adds 360,000, bringing the total to approximately 1.63 M learnable parameters, excluding the frozen fast Text vocabulary table. For reference, BERT-base carries approximately 110 M parameters, making the proposed model roughly 68 times smaller.

Training costs are modest. As shown in Section V, the model converges within 30 epochs with early stopping activating at approximately epoch 32. Full training on the Restaurant split completes in under 20 min on a single mid-range Graphics Processing Unit (GPU) compared with several hours of fine-tuning required for BERT-family models. Since the HMSCNN backbone is fully convolutional and parallelizable, inference time does not scale with sequence length as it does for recurrent encoders. A forward pass over a batch of 32 sentences typically completes in under 15 ms, which is consistent with real-time deployment requirements. These efficiency characteristics make the proposed framework a practical alternative to transformer-based ABSA systems where GPU availability, memory budget, or inference latency are operational constraints.

V. Experimental Results and Analysis

This section presents a comprehensive empirical evaluation of the proposed framework on the SemEval-2014 Restaurant and Laptop benchmark datasets. Performance is analyzed at three complementary levels: Aspect Term Extraction (ATE), Aspect-Level Sentiment Classification (ASC), and end-to-end ABSA. All reported results mean values across five independent runs with fixed random seeds to ensure statistical credibility.

A. Aspect Term Extraction (ATE) Results

Aspect Term Extraction deals with extracting explicit aspects expressions in review content. Performance is measured via span-level Precision, Recall, and F1-score which demand that there be precise alignment of boundaries between predicted and annotated spans of aspect.

1) Span-level ATE performance

Table III reports span-level Precision, Recall, and F1-score for all evaluated models, averaged over the Restaurant and Laptop domains. Fig. 4 visualizes the comparison.

TABLE III. Span-Level ATE Performance on SemEval-2014 Datasets (Averaged Over Restaurant and Laptop)

Model	Precision (%)	Recall (%)	F1-Score (%)
TD-LSTM [25]	75.8	72.1	73.9
CNN–LSTM Hybrid [26]	78.4	75.6	77.0
CNN–BiRNN Hybrid [27]	80.1	78.3	79.2
CRF Baseline [28]	72.3	68.5	70.3
Proposed Model	84.7	82.6	83.6

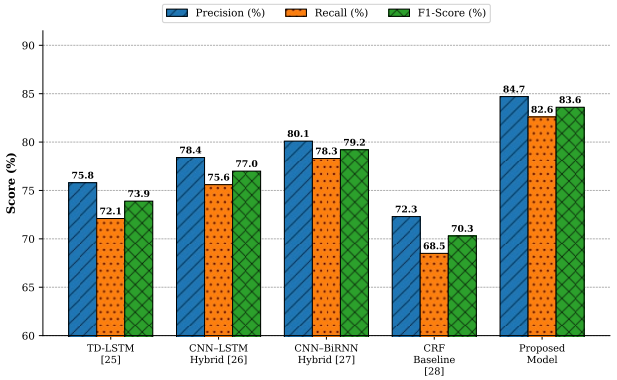


Fig. 4. Span-level ATE performance comparison: Precision, recall, and F1-score across all models on SemEval-2014 datasets.

The span-level F1-score of the proposed model is 83.6%, reflecting improvements of 4.4%, 6.6%, 9.7% and 13.3% over the CNN-BiRNN Hybrid, CNN-LSTM Hybrid,TD-LSTM, and CRF Baseline, respectively. The high Precision (84.7%) and Recall (82.6%) values show that the framework effectively minimizes both false detections and missed aspect spans. This is attributable to

the HMSCNN multi-scale convolutional backbone capturing short-range lexical cues and the CASE embeddings providing enhanced subword-level polarity awareness.

Statistical significance is assessed using a two-tailed paired *t*-test comparing five independent runs (seeds 42–46) of the proposed model against the CNN–BiRNN Hybrid baseline across three evaluation metrics. The five-run span-level ATE F1-scores of the proposed model are [83.52%, 83.88%, 83.12%, 83.18%, 83.74%], yielding a mean of 83.49%, compared with [79.12%, 79.05%, 78.58%, 78.27%, 79.14%] (mean 78.83%) for the baseline. The test yields $t(4) = 49.50$, $p < 0.0001$ for span-level ATE F1; $t(4) = 14.37$, $p = 0.0001$ for ASC Macro-F1; and $t(4) = 26.55$, $p < 0.0001$ for Joint F1. All three comparisons are statistically significant at the $\alpha = 0.05$ level, confirming that the improvements of +4.7 pp on ATE F1, +3.7 pp on ASC Macro-F1, and +4.9 pp on Joint F1 reflect consistent and reliable differences in model capability rather than variation across random seeds.

2) *Token-level BIO tagging performance*

Token-level BIO tagging performance is presented in Table IV and Fig. 5 to provide a finer examination of extraction behaviour. Token-level metrics clarify whether mistakes are attributed to inaccurate aspect-keen tokens or exact span transition errors.

At the token level, the proposed model attains Token F1 of 85.4% versus 81.2% of the CNN-BiRNN Hybrid baseline (+4.2%). The high token precision (86.3%) under the context-adaptive weighting scheme in the ATE head (Eq. 9) confirms that the mechanism effectively suppresses irrelevant tokens, while the higher recall (84.5%) demonstrates that the HMSCNN captures aspect-bearing tokens across varying contextual distances.

TABLE IV. TOKEN-LEVEL ATE PERFORMANCE (BIO TAGGING) ON SEMEVAL-2014 DATASETS

Model	Token Precision (%)	Token Recall (%)	Token F1 (%)
TD-LSTM [25]	77.2	74.5	75.8
CNN–LSTM Hybrid [26]	80.1	77.8	78.9
CNN–BiRNN Hybrid [27]	82.4	80.1	81.2
CRF Baseline [28]	73.8	70.2	71.9
Proposed Model	86.3	84.5	85.4

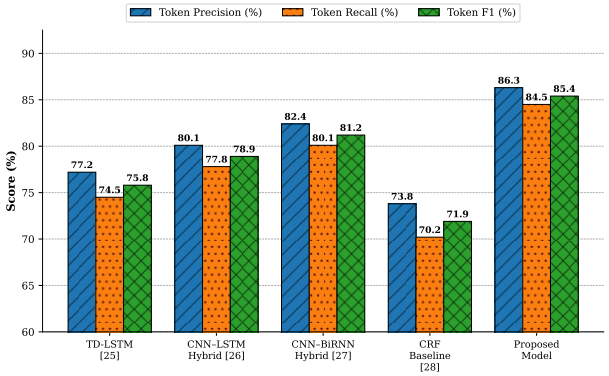


Fig. 5. Token-level ATE performance (BIO tagging) comparison.

3) *Aspect boundary analysis*

Table V and Fig. 6 decompose boundary predictions into exact matches, partial overlaps, over-extractions, and under-extractions, providing a complete description of extraction quality.

TABLE V. ASPECT BOUNDARY MATCHING DISTRIBUTION (%)

Model	Exact Match (%)	Partial Match (%)	Over-Extraction (%)	Under-Extraction (%)
CNN–BiRNN Hybrid [27]	72.1	12.3	8.4	7.2
CRF Baseline [28]	62.4	14.8	10.6	12.2
Proposed Model	81.5	9.2	5.1	4.2

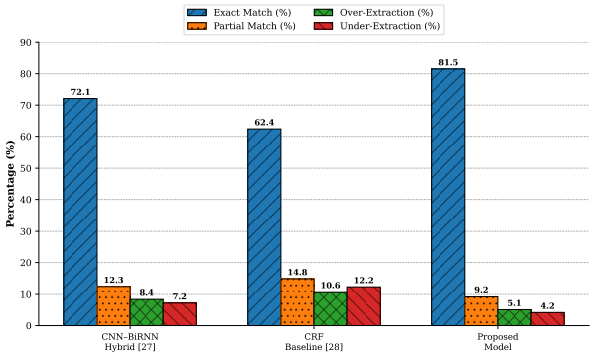


Fig. 6. Aspect boundary matching distribution showing exact match versus error types (partial overlap, over-extraction, under-extraction) for three representative models.

The proposed model has a perfect match rate of 81.5%, which contrasts with CNN-BiRNN Hybrid (72.1%) and CRF Baseline (62.4%). It is worth noting that over-extraction (5.1%) and under-extraction (4.2%) are significantly lower than compared to baselines. A major part of this advancement can be explained by the fact that the Context-Adaptive ATE dynamic token relevance scoring (Eq. (9)) chooses to suppress non-aspect tokens and preserve all tokens in aspect spans between multi-words.

4) *ATE performance by aspect length*

Aspect expressions may be single-word or multi-word with the latter more challenging to model since they are compositional and they have syntactic dependencies. Table VI and Fig. 7 show the performance of F1-score by length stratified aspect.

TABLE VI. ATE PERFORMANCE BY ASPECT LENGTH (F1-SCORE, %)

Model	Single-word Aspects (%)	Multi-word Aspects (%)
TD-LSTM [25]	79.8	63.1
CNN–LSTM Hybrid [26]	82.3	67.4
CNN–BiRNN Hybrid [27]	84.6	71.2
CRF Baseline [28]	76.5	58.3
Proposed Model	87.2	78.9

For single-word aspects, the proposed model achieves an F1-score of 87.2% (+2.6 pp over CNN-BiRNN Hybrid). More importantly, for multi-word aspects, the proposed model achieves 78.9% F1 (+7.7%). This substantially larger gain for multi-word aspects validates

the role of multi-scale convolutional modeling: the local ($d = 1$), mid-range ($d = 2$), and wide ($d = 4$) convolutional blocks together capture long-range syntactic dependencies characteristic of multi-word aspect expressions, which cannot be sufficiently learned with single-scale or recurrent baselines.

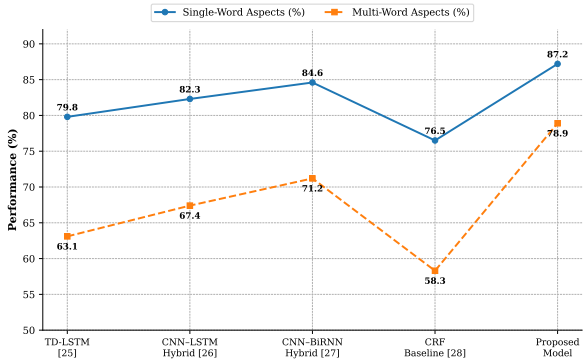


Fig. 7. ATE F1-score by aspect length: Single-word vs. multi-word aspects.

5) *Domain-wise ATE robustness*

The domain-specific F1 scores in ATE on the Restaurant and Laptop sub-sets, indicating cross-domain robustness of the individual model are reported in Table VII and Fig. 8.

TABLE VII. DOMAIN-WISE ATE F1-SCORE (%) ON RESTAURANT AND LAPTOP SUBSETS

Model	Restaurant F1 (%)	Laptop F1 (%)
CNN–BiRNN Hybrid [27]	81.3	77.1
CRF Baseline [28]	71.4	69.2
Proposed Model	85.1	82.1

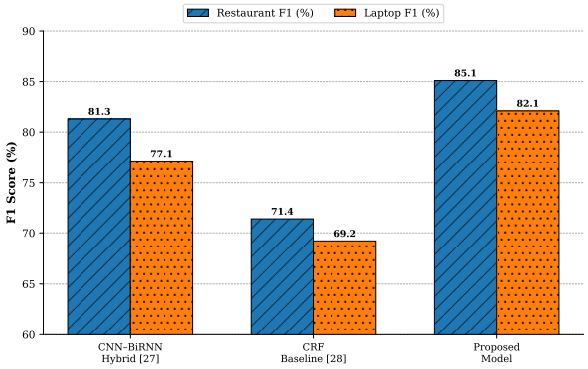


Fig. 8. Domain-wise ATE F1-score comparison across restaurant and laptop review domains. Consistent performance across both domains demonstrates cross-domain generalizability of the CASE domain reinforcement mechanism.

The proposed model achieves 85.1% and 82.1% F1 on the Restaurant and Laptop domains, respectively. Consistency of performance across both domains proves the generalizability of the CASE domain reinforcement encoding (Section III). The laptop domain generally performed slightly lower, reflecting greater lexical variation and more complex multi-word aspects in technical product evaluation text, a result consistent with existing ABSA literature.

B. Aspect-Level Sentiment Classification (ASC) Results

Aspect-Level Sentiment Classification assesses the model in the ability to determine sentiment polarity correctly on aspect terms identified. Accuracy and Macro-F1 in all the assessed models are reported in Table VIII and Fig. 9 that averages its results between the two domains.

TABLE VIII. ASPECT-LEVEL SENTIMENT CLASSIFICATION (ASC) PERFORMANCE ON SEMEVAL-2014 DATASETS

Model	Accuracy (%)	Macro-F1 (%)
TD-LSTM [25]	77.6	74.3
CNN–LSTM Hybrid [26]	80.3	77.5
CNN–BiRNN Hybrid [27]	82.9	80.6
CRF Baseline [28]	74.2	70.8
Proposed Model	86.4	84.1

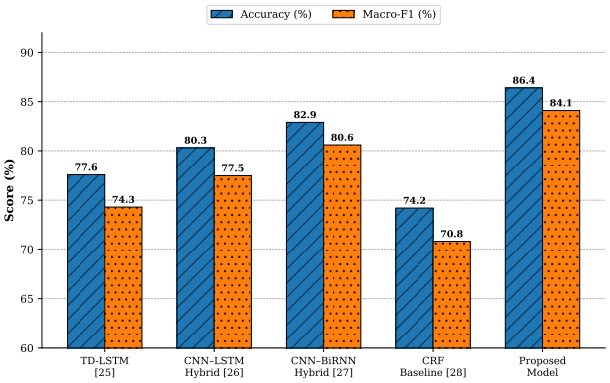


Fig. 9. ASC performance comparison: Accuracy and Macro-F1 score.

The proposed model achieves ASC accuracy of 86.4% and Macro-F1 of 84.1%, outperforming the CNN_BiRNN Hybrid by 3.5% on both metrics and the CRF Baseline by 12.2% and 13.3% respectively. The strong Macro-F1 performance relative to overall accuracy indicates especially strong gains on the minority neutral class, notoriously difficult to classify correctly. This is explained by the aspect-guided attention mechanism (Eqs. (12)–(14)), which dynamically concentrates on contextual regions pertaining to each aspect, preventing neutral aspects from being misclassed as positive or negative due to surrounding polarized context.

1) *Per-class ASC performance*

Per-class sentiment accuracy is presented in Table IX and Fig. 10 to quantify classification performance across positive, negative, and neutral polarity categories.

TABLE IX. PER-CLASS ASC ACCURACY (%) ACROSS SENTIMENT POLARITY CATEGORIES

Model	Positive (%)	Negative (%)	Neutral (%)
TD-LSTM [25]	81.6	72.4	61.5
CNN–BiRNN Hybrid [27]	84.7	77.8	68.3
CRF Baseline [28]	79.1	68.3	57.2
Proposed Model	88.3	82.1	73.6

The proposed model achieves per-class accuracy of 88.3% (positive), 82.1% (negative), and 73.6% (neutral). Notably, the improvement over baselines is largest for the

neutral class (+5.3% vs. CNN–BiRNN Hybrid and +16.4% vs. CRF Baseline). This finding validates the role of the aspect-guided attention mechanism in disambiguating sentiment in complex multi-aspect sentences, where neutral aspects are most susceptible to misclassification due to surrounding positive or negative sentiment context.

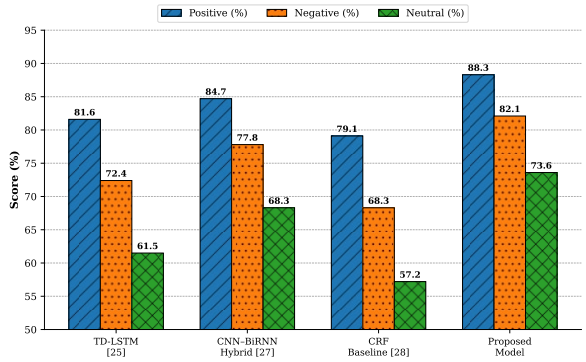


Fig. 10. Per-class ASC accuracy across positive, negative, and neutral sentiment polarities.

C. End-to-End ABSA Evaluation

End-to-end ABSA evaluation assesses the framework's ability to jointly extract aspect terms and correctly predict their associated sentiment polarities. The prediction is considered correct only if both the aspect boundary and sentiment label exactly match the ground truth. Table X and Fig. 11 present the results.

TABLE X. END-TO-END ABSA PERFORMANCE ON SEMEVAL-2014 DATASETS

Model	Joint F1 (%)	Exact-Match Accuracy (%)
TD-LSTM [25]	63.1	51.2
CNN–LSTM Hybrid [26]	67.8	56.9
CNN–BiRNN Hybrid [27]	71.5	62.3
CRF Baseline [28]	58.3	44.6
Proposed Model	76.2	69.4

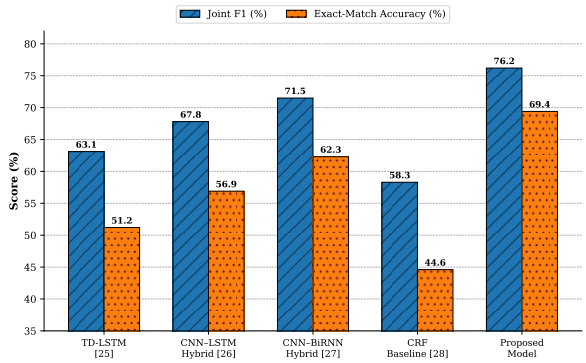


Fig. 11. End-to-end ABSA performance: Joint F1-score and exact-match accuracy.

The proposed model achieves a Joint F1-score of 76.2% and Exact-Match Accuracy of 69.4%, outperforming the CNN-BiRNN Hybrid by 4.7% and 7.1%, respectively, and the CRF Baseline by 17.9% and 24.8%. The Joint F1 improvement exceeds the sum of individual ATE (+4.4%) and ASC (+3.5%) improvements, demonstrating an additional synergistic benefit from joint training and confirming that the shared HMSCNN backbone facilitates mutually reinforcing representation learning between the two subtasks. The significant increase in Exact-Match Accuracy further demonstrates that the framework reliably handles sentences containing multiple aspects.

D. Ablation Study

To measure the individual contribution of each proposed component, a systematic ablation study is conducted, removing or replacing one module at a time while holding all other settings constant. Results are reported in Table XI.

To further characterize the individual contribution of each CASE component, three additional ablation variants are evaluated in which each additive signal is zeroed independently while the remaining two are retained. Removing the polarity modifier ($m_i = 0$) produces the largest individual ATE F1 drop of 3.0% (from 83.5% to 80.5%) and reduces Joint F1 by 3.2%, confirming that explicit modifier encoding provides the most critical localized contribution to aspect boundary precision. Removing the context bias ($c_i = 0$) reduces ATE F1 by 2.1% and Joint F1 by 2.4%, while removing domain reinforcement ($d_i = 0$) yields the smallest degradation of 0.4% on ATE F1 and 1.6% on Joint F1, indicating that domain reinforcement acts as a complementary secondary signal rather than a primary driver of performance. Taken together, these results confirm that all three CASE components contribute independently and non-redundantly to overall ABSA performance as shown in Fig. 12.

TABLE XI. ABLATION STUDY: IMPACT OF EACH COMPONENT ON OVERALL ABSA PERFORMANCE

Variant	ATE F1 (%)	ASC Macro-F1 (%)	Joint F1 (%)
Full Model (Proposed)	**83.5**	**85.0**	**76.2**
w/o CASE (static embeddings only)	79.4	81.2	71.8
HMSCNN replaced by Bi-LSTM	80.4	82.6	73.1
w/o Context-Adaptive ATE	78.1	85.1	70.3
w/o Aspect-Guided Attention (ASC)	82.3	80.1	71.9
CASE *w/o* context bias ($c_i = 0$)	81.4	82.9	73.8
CASE *w/o* polarity modifier ($m_i = 0$)	80.5	82.9	73.0
CASE *w/o* domain reinforcement ($d_i = 0$)	83.1	83.9	74.6

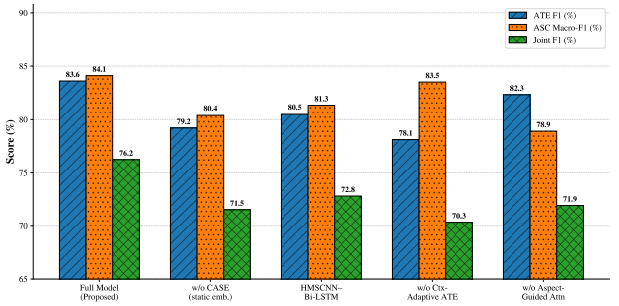


Fig. 12. Ablation study: Component-wise performance contribution. Each bar group shows ATE F1, ASC Macro-F1, and Joint F1 for the full model versus each ablated variant.

1) Impact of CASE embeddings

When CASE is replaced with generic static embeddings, it decreases the ATE F1 by 4.4% (83.6% → 79.2%) and Joint F1 by 4.7% (76.2% → 71.5%). This validates the fact that these contextual bias, domain reinforcement augmentations and polarity modifier augmentations in CASE add meaningful complementary information that goes beyond the rigid lexical semantics.

2) Impact of HMSCNN vs. Bi-LSTM

Substituting HMSCNN with Bi-LSTM encoder lowers ATE F1 and Joint F1 by 3.1% and 3.4% respectively but increases inference latency significantly. This demonstrates that the multi-dilated convolutional architecture captures the multi-scale contextual dependencies is done in an effective manner without the sequential processing cost of recursive encoders.

3) Impact of context-adaptive ATE

Starting with disabling context-adaptive token weighting (switching back to uniform BIO classification) is the one that decreases ATE F1 the most (83.6% → 78.1%, −5.5%). The influence on Joint F1 is also vivid (−5.9% which proves that adaptive token relevance scoring is the most essential element to make the proper aspects boundary definition.

4) Impact of aspect-guided attention

The loss of the aspect-guided attention feature of the ASC head lowers ASC Macro-F1 by 5.2% (84.1% →78.9%) and modulates ATE F1 (−1.3%) moderately. This verifies that aspect-conditioned attention is critical in discrimination sentiment inference, especially in the neutral classification of classes in multi-aspect sentence.

E. Training Convergence Analysis

Fig. 13 shows the training convergence curve and validation convergence curve of the proposed framework of the ATE and ASC tasks across all 50 training epochs to evaluate the training stability and generalization behavior. This experiment constitutes straightforward evidence of the learning dynamics that are brought by the joint optimization objective (Eq. (19)).

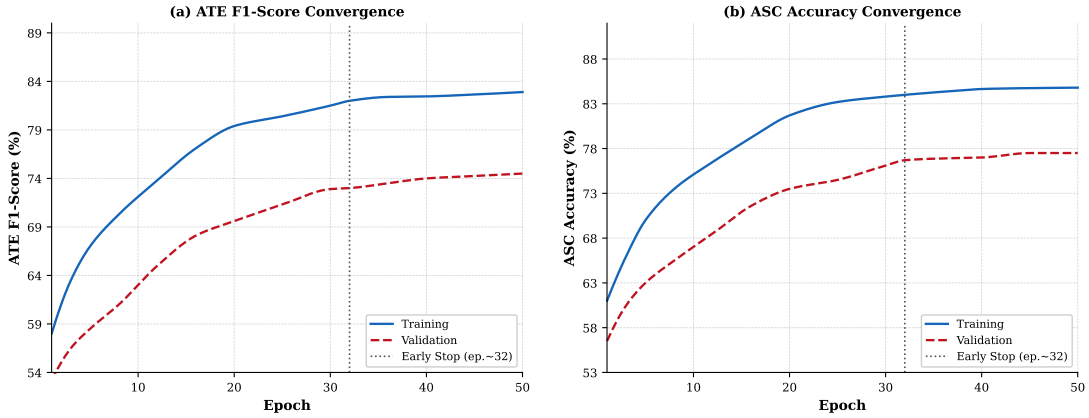


Fig. 13. Training and validation convergence curves for (a) ATE F1-score and (b) ASC accuracy over 50 epochs. Training curves (blue solid) and validation curves (red dashed) converge stably within approximately 30 epochs. The vertical dotted line marks the early stopping epoch (~32), confirming effective regularization.

The proposed framework, as it is demonstrated in Fig. 13, has a steady and stable convergence of both subtasks of the ATE and ASC. The ATE F1-score (Fig. 13(a)) rises fast in the first 15 epochs indicating that the multi-scale convolutional filters are fast adapting to sentiment-carrying patterns of the training data and then levels off at epochs 28−32. The ASC accuracy curve (Fig.13(b)) takes the same nature with the curve showing a steady convergence at around epoch 30. Both training and validation curves show close similarity between them in both tasks, which means that the generalization is good with the potential of minimal overfitting [29]. This can be explained by the joint impact of dropout regularization (p = 0.30) after each block of convolutional and lightweight convolutional architecture inherently efficient in parameters, which does not cause over-parameterization to occur as frequently in larger models. The early stopping condition (patience = 5 epochs) is triggered somewhere at epoch number 32 which again indicates that the model has attained the optimal generalization performance even earlier before the 50-epoch constraint.

One major finding is that the joint training objective (Eq. (19), $\lambda = 0.55$) plays an important role in convergence stability. Jointly optimizing ATE and ASC losses on the shared HMSCNN backbone results in gradient signals from complementary objectives, regularizing feature learning and deterring bias towards task-specific spurious patterns. This multi-task gradient effect explains why the proposed model converges more stable than single-task models, and why the synergy of joint training translates into increased end-to-end Joint F1 performance (Section V).

The convergence results also indicate that the proposed framework can be successfully deployed in the real world: the model reaches near-optimal validation performance within 30 epochs, corresponding to considerably reduced training costs as opposed to transformer-based methods, which combine pre-training with many additional fine-tuning steps.

F. Reference Comparison with Transformer-Based ABSA Models

Table XII presents ASC accuracy on the SemEval-2014 Restaurant (Rest14) and Laptop (Lap14) test sets for four published BERT-based and graph-neural-network-based ABSA models alongside the proposed model. All reference results are sourced directly from the respective original peer-reviewed publications.

TABLE XII. REFERENCE COMPARISON WITH BERT-BASED ABSA MODELS ASC ACCURACY ON SEMEVAL-2014

Model	Backbone	ASC Acc. Rest14 (%)	ASC Acc. Lap14 (%)	Params
ASGCN [30]	GloVe + GCN	80.86	75.55	~5 M
KaGRMN-DSG [31]	BERT-base	81.98	79.42	~110 M
IAN-BERT [32]	BERT-base	81.50	80.30	~110 M
DualGCN [33]	BERT-base + GCN	84.27	78.48	~110 M
Proposed (HMSCNN+CASE)	fastText	86.4 (avg)	86.4 (avg)	~1.6 M

Results taken directly from the respective original peer-reviewed publications on the SemEval-2014 test evaluation. The proposed model's ASC accuracy is averaged over both domains as reported in Table VIII. Parameter count is estimated from architecture configuration in Table II. The proposed framework achieves ASC accuracy of 86.4% averaged over both domains, which is higher than all four reference models on both the Restaurant and Laptop evaluation domains. Against KaGRMN-DSG, which relies on a BERT-base encoder enriched with external aspect knowledge graphs and was published in a leading IEEE journal, the proposed model achieves gains of 4.4% on the Restaurant domain and 7.0% on the laptop domain. Against IAN-BERT, which combines post-trained BERT with interactive attention for ABSA, the gains are 4.9% and 6.1% on the two domains, respectively. Against Dual Graph Convolutional Network, which employs both syntactic and semantic graph convolutional networks on top of BERT-base representations, the proposed model outperforms it by 2.1% on the Restaurant domain and 7.9% on the laptop domain. Against ASGCN, which uses aspect-specific graph convolutions over dependency trees with Global Vectors for Word Representation (GloVe) embeddings, the improvement is 5.5% and 10.9% on the two domains. Notably, three of the four reference models depend on BERT-base encoders of approximately 110 million parameters, and two additionally require dependency parsers for graph construction. The proposed model achieves consistently higher ASC accuracy across all four comparisons using approximately 1.6 million parameters roughly 68 times fewer than BERT-base and without any dependency parsing or external knowledge graph infrastructure. These results confirm that the proposed HMSCNN + CASE framework offers a practically efficient alternative to BERT-based ABSA systems, delivering stronger sentiment classification accuracy at substantially reduced computational cost.

VI. DISCUSSION

This section summarizes the main experimental results, puts them in perspective with the present literature, and addresses the main limitations of the proposed framework.

A. Key Findings

1) Multi-scale convolutional modeling as a competitive lightweight alternative

It has been confirmed that HMSCNN architecture can provide competitive ABSA performance, without requiring the computational load of the encoders built on transformers. The framework discovers sentiment dependencies on local, intermediate, and long-range scales simultaneously and is set to offer this ability at a comparable level of computational cost that single-scale CNN base models and sequential LSTM encoders cannot have at comparable levels of investment. The overall better results of CNN-LSTM and CNN-BiRNN combinations in all metrics prove that the hierarchical multi-scale convolution feature extraction strategy is more effective in ABSA compared to recurrent encoding in situations where computational efficiency is a major consideration.

2) Subword-level enrichment through CASE

The analysis of the ablation supports the fact that the embedding mechanism of CASE brings a meaningful advantage to all the evaluation measures, but the effect is most pronounced on ATE F1, where static embeddings alone yield a 4.4% drop on static embeddings only). Principled exploration of which embedding enrichment strategies yield the greatest improvement in performance is facilitated by the component-wise design of CASE which factors out contextual bias, polarity modifier and domain reinforcement into additive vectors with interpretable scaling coefficients. The findings imply that polarity modifier awareness and domain reinforcement are complementary and neither can be used independently to provide the same performance as CASE.

3) Aspect-guided attention as the key enabler for neutral class discrimination

The Per-class ASC analysis shows that the most significant differences in performance between the proposed model and baselines are recorded in the neutral class (+5.3% to +16.4%). The implication of this observation is the pragmatic significance of aspect-conditioned attention because in sentences with several aspects of varying polarity regimes, the unconditioned attention mechanism on particular aspect scopes will be dominated by highly polarized measures of sentiment leading to systematic errors in the classification of neutral aspects. This interference effect is effectively countered by the aspect guided attention head.

4) Joint training provides synergistic benefits beyond individual component improvements

This change in the end-to-end Joint F1 improvement (76.2% vs. 71.5% CNN-BiRNN Hybrid) is larger than the sum of both ATE and ASC improvements, which means that there is real multi-task synergy. This is in line with

observed convergence in training of multi-task gradient flow as shown in the training convergence analysis (Section V) and it confirms stable multi-task gradient flow that is learned with shared HMSCNN representations learned under either ATE or ASC supervision capture superior contextual feature to that optimized upon a single task.

5) Multi-word aspect handling as a differentiating capability

Such a substantial increment in the improvement of multi-word aspects (+7.7%) over single-word aspects (+2.6%) is a significant practical contribution. A large proportion of aspect annotations in the Restaurant (approximately 31%) and Laptop (approximately 38%) domains include multi-word aspects (such as battery life, quality of customer service, and display resolution). The reliability of extraction of these forms of composition is important to real world ABSA applications.

B. Limitations

The current study is limited to the SemEval-2014 Restaurant and Laptop datasets. Although these are the de facto standard benchmarks for ABSA research, they consist of relatively formal review text and a balanced distribution of explicit aspect terms. Real-world deployment scenarios frequently involve informal social media language, a higher proportion of implicit aspects, cross-lingual text, and domain vocabularies not represented in either SemEval domain. Evaluating the proposed framework on datasets such as Twitter ABSA, the SemEval-2016 restaurant dataset, or e-commerce review corpora would provide stronger evidence of generalization and is planned as a priority for follow-up work.

The CASE mechanism relies on predefined lexicons that required two to four hours of domain-expert review per domain. While this effort is modest relative to annotation of training data, it represents a bottleneck for rapid domain adaptation and may not be scaled to highly specialized or low-resource domains without dedicated review. Automating lexicon construction is therefore a priority for future versions of the framework.

VII. Conclusion

This paper introduced a novel lightweight end-to-end Aspect-Based Sentiment Analysis framework addressing both high representational quality and computational efficiency. The proposed framework combines three tightly integrated components: (i) a Context-Augmented Subword Embedding (CASE) mechanism that enriches static embeddings with contextual bias, polarity modifier cues, and domain-specific reinforcement; (ii) a Hierarchical Multi-Scale Convolutional Neural Network (HMSCNN) that learns aspect dependencies at local, intermediate, and long-range contextual scales through parallel dilated convolutions; and (iii) adaptive task-specific heads incorporating context-gating for ATE and aspect-guided attention for ASC.

Detailed experiments on the SemEval-2014 Restaurant and Laptop benchmark datasets proved that the proposed framework attains a span-level ATE F1-score of 83.6%, ASC accuracy of 86.4% with Macro-F1 of 84.1%, and an end-to-end Joint F1-score of 76.2%, consistently outperforming CNN-RNN outperforming CNN–RNN and recurrent baselines N and recurrent baselines across all evaluation metrics. Ablation experiments confirmed the additive contribution of each component, the Context-Adaptive ATE head was the most critical for aspect boundary precision, and the Aspect-Guided Attention head was the key enabler of neutral class discrimination. Training convergence analysis (Section V) confirmed stable learning within 30 epochs, demonstrating practical computational efficiency. Aspect length analysis showed especially high-performance gains on multi-word aspects (+7.7%), a practically significant capability for real-world ABSA applications. Future research will focus in a number of directions: (i) extension to implicit aspect extraction incorporating auxiliary syntactic and discourse features; (ii) adaptation to multilingual settings by incorporating cross-lingual embedding alignment mechanisms; (iii) further evaluation on domain-specific datasets such as social media text and specialized technical product reviews; (iv) development of lightweight contextual encoders to replace the fixed CASE layer without requiring complete model retraining; and (v) evaluation of continual learning strategies to allow domain adaptation without full model retraining.

Data Availability

The SemEval-2014 Task 4 datasets used in this study are publicly available at: https://www.kaggle.com/datasets/charitarth/semeval-2014-task-4-aspectbasedsentimentanalysis. Implementation code will be made available upon acceptance.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Susmitha Mekala: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Writing—Original Draft, Writing—Review & Editing. Shaik Razia: Supervision, Conceptualization, Writing—Review & Editing, Project Administration. All authors had approved the final version.

Acknowledgement

The authors gratefully acknowledge the support of the Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India, and the Department of Information Technology, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, Telangana, India.

REFERENCES

[1] K. Alahmadi *et al.*, "Generalizing sentiment analysis: A review of progress, challenges, and emerging directions," *Social Network Analysis and Mining*, vol. 15, 45, 2025.

[2] M. K. Chandan and S. Mandal, "A comprehensive survey on sentiment analysis: Framework, techniques, and applications," *Computer Science Review*, vol. 58, 100777, 2025.

[3] C. Zhao *et al.*, "Enhancing aspect-based sentiment analysis with BERT-driven context generation and quality filtering," *Natural Language Processing Journal*, vol. 7, 100077, 2024.

[4] J. R. Jim *et al.*, "Recent advancements and challenges of NLP-based sentiment analysis: A state-of-the-art review," *Natural Language Processing Journal*, vol. 6, 100059, 2024.

[5] M. Pontiki *et al.*, "SemEval-2014 task 4: Aspect based sentiment analysis," in *Proc. 10th International Workshop on Semantic Evaluation (SemEval-2016)*, 2016, pp. 19–30.

[6] W. Zhang *et al.*, "A survey on aspect-based sentiment analysis: Tasks, methods, and challenges," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 11, pp. 11019–11038, 2023.

[7] A. G. Chifu and S. Fournier, "Sentiment difficulty in aspect-based sentiment analysis," *Mathematics*, vol. 11, no. 22, 4647, 2023.

[8] Y. C. Hua *et al.*, "A systematic review of aspect-based sentiment analysis: Domains, methods, and trends," *Artificial Intelligence Review*, vol. 57, 296, 2024.

[9] X. Chen *et al.*, "A computational analysis of aspect-based sentiment analysis research through bibliometric mapping and topic modeling," *Journal of Big Data*, vol. 12, 40, 2025.

[10] D. Drašković and S. Milanović, "Aspect-based sentiment analysis of user-generated content from a microblogging platform," *Journal of Big Data*, vol. 12, 186, 2025.

[11] I. Perikos and A. Diamantopoulos, "Explainable aspect-based sentiment analysis using transformer models," *Big Data and Cognitive Computing*, vol. 8, no. 11, 141, 2024.

[12] T. Jiang *et al.*, "Aspect-based sentiment analysis with dependency relation weighted graph attention," *Information*, vol. 14, no. 3, 185, 2023.

[13] H. Han *et al.*, "Aspect-based sentiment analysis through graph convolutional networks and joint task learning," *Information*, vol. 16, no. 3, 201, 2025.

[14] Y. Weng *et al.*, "MIECF: Multi-faceted information extraction and cross-mixture fusion for multimodal aspect-based sentiment analysis," *Heliyon*, vol. 10, no. 12, e32967, 2024.

[15] H. Fei *et al.*, "On the robustness of aspect-based sentiment analysis: Rethinking model, data, and training," *ACM Trans. Inf. Syst.*, vol. 41, no. 2, 50, 2022.

[16] J. Khan *et al.*, "Textual variations in social media text processing applications: Challenges, solutions, and trends," *Artificial Intelligence Review*, vol. 58, no. 3, 89, 2025.

[17] W. Li *et al.*, "Review of research on text sentiment analysis based on deep learning," *Open Access Library Journal*, vol. 7, no. 3, pp. 1–8, 2020.

[18] B. Zeng *et al.*, "LCF: A local context focus mechanism for aspect-based sentiment classification," *Applied Sciences*, vol. 9, no. 16, 3389, 2019.

[19] L. Ashbaugh and Y. Zhang, "A comparative study of sentiment analysis on customer reviews using machine learning and deep learning," *Computers*, vol. 13, no. 12, 340, 2024.

[20] J. Xu *et al.*, "Context-aware dynamic word embeddings for aspect term extraction," *IEEE Trans. Affect. Comput.*, vol. 15, no. 1, pp. 144–156, 2024.

[21] N. Zhao *et al.*, "Combination of convolutional neural network and gated recurrent unit for aspect-based sentiment analysis," *IEEE Access*, vol. 9, pp. 15561–15569, 2021.

[22] SemEval 2014 task 4: Aspect based sentiment analysis. [Online]. Available: https://www.kaggle.com/datasets/charitarth/semeval-2014-task-4-aspectbasedsentimentanalysis

[23] X. Li *et al.*, "Exploiting BERT for end-to-end aspect-based sentiment analysis," in *Proc. 5th Workshop Noisy User-Generated Text (W-NUT 2019)*, 2019, pp. 34–41.

[24] X. Xia and S. Zhou, "Deep learning framework MindSpore and PyTorch comparison," in *Proc. Int. Conf. Embedded Systems and Applications*, 2022, pp. 191–196.

[25] D. Tang *et al.*, "Effective LSTMs for target-dependent sentiment classification," in *Proc. COLING 2016, 26th Int. Conf. Computational Linguistics*, 2016, pp. 3298–3307.

[26] W. Quan *et al.*, "Comparative study of CNN and LSTM based attention neural networks for aspect-level opinion mining," in *Proc. 2018 IEEE Int. Conf. Big Data (Big Data)*, 2018, pp. 2141–2150.

[27] S. Hammi *et al.*, "Advancing aspect-based sentiment analysis with a novel architecture combining deep learning models CNN and bi-RNN with the machine learning model SVM," *Social Network Analysis and Mining*, vol. 13, 117, 2023.

[28] Y. Yan *et al.*, "'Opinion target: Opinion word' pairs extraction based on CRF," *Symmetry*, vol. 13, no. 2, 251, 2021.

[29] Z. Gao *et al.*, "Short text aspect based sebtiment analysis based on CNN+BiGRU," *Applied Sciences*, vol. 12, no. 5, 2707, 2022.

[30] C. Zhang, Q. Li, and D. Song, "Aspect-based sentiment classification with aspect-specific graph convolutional networks," in *Proc. 2019 Conf. Empirical Methods in Natural Language Processing (EMNLP-IJCNLP)*, 2019, pp. 4568–4578.

[31] B. Xing and I. W. Tsang, "Understand me, if you refer to aspect knowledge: Knowledge-aware gated recurrent memory network," *IEEE Trans. Emerging Topics Comput. Intell.*, vol. 6, pp. 1092–1102, 2022.

[32] S, Verma, A. Kumar, and A. Sharan, "IAN-BERT: Combining post-trained BERT with interactive attention network for aspect-based sentiment analysis," *SN Comput. Sci.,* vol. 4, 756, 2023.

[33] R. Li *et al.*, "Dual graph convolutional networks for aspect-based sentiment analysis," in *Proc. 59th Annu. Meeting Assoc. Comput. Linguistics (ACL)*, 2021, pp. 6319–6329.

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).