Journal of Computer Technology and Software

ISSN:2998-2383

Vol. 4, No. 5, 2025

Perception-Guided Structural Framework for Large Language Model Design

Fan Guo[1], Lin Zhu[2], Yichen Wang[3], Guohui Cai[4]
[1]Illinois Institute of Technology, Chicago, USA
[2]Stevens Institute of Technology, New Jersey, USA
[3]Georgia Institute of Technology, Atlanta, USA
[4]Illinois Institute of Technology, Chicago, USA
*Corresponding Author: Guohui Cai; gcai3@hawk.iit.edu

Abstract: This paper proposes a structural optimization method for large language models based on a perception-representation integration mechanism. The goal is to enhance semantic construction and contextual consistency under complex language scenarios. The method introduces perceptual feature extraction modules and perception-guided attention mechanisms. This enables dynamic semantic modeling of language input and multi-level structural-perception interaction. It addresses the disconnect between representation and structure in traditional language models. In implementation, the method integrates a perception-driven representation update strategy into the GPT architecture. It constructs a perception graph to regulate attention distribution. This design improves the model's structural expressiveness. Experiments on the WikiText-103 dataset show that the proposed method outperforms mainstream language models in key metrics, including Perplexity, BLEU, and Semantic Consistency. Additionally, a series of hyperparameter sensitivity experiments and comparative analyses of perception injection strategies are conducted. These evaluate the impact of structural components on model performance. The results confirm the stability and effectiveness of the proposed mechanism under different training configurations.

Keywords: Large language model, perception mechanism, representation modeling, attention optimization

1. Introduction

In recent years, artificial intelligence has developed rapidly. Large language models, as a key support technology in natural language processing, have shown outstanding performance in tasks such as text generation, language understanding, and logical reasoning. In particular, Transformer-based pre-trained models have expanded the frontiers of general language intelligence. They do so by leveraging strong expressive power and transferability. However, with the exponential growth in model size, significant structural bottlenecks remain. These include understanding complex contexts, maintaining semantic consistency, and improving cognitive depth. This suggests that despite their engineering success, current models still lack a unified theoretical framework for integrating perception and representation in semantic modeling, context preservation, and pragmatic inference[1].

Traditional language models rely heavily on static embeddings and hierarchical mappings. They use stacked attention modules to capture and abstract semantic features. While such architectures offer some generalization at the surface level, they lack the ability to dynamically deconstruct and reconstruct input data. As a result, they fail to integrate internal conceptual relationships at higher semantic levels. This structural disconnection limits their ability to model complex language phenomena[2]. Examples include metaphor, contextual association, and cross-paragraph memory. The root cause lies in the absence of an integrated mechanism linking language perception with semantic representation. In human cognition, language generation and understanding are not linear processes. They emerge from complex interactions between sensory inputs and representational systems. This insight prompts a reevaluation of language model design. A new modeling path is needed-one that deeply fuses perception with representation[3].

The idea of integrating perception and representation originates from cognitive science, neurolinguistics, and complex systems theory. These fields suggest that language formation is not a one-way process. It involves multi-level feedback among perception, structural encoding, and high-level abstraction. From this view, large language models should evolve from passive mappers to active builders of semantic fields. This shift helps improve semantic consistency and expressive completeness. It also provides new theoretical support for addressing challenges such as context integration, cross-sentence inference, and semantic stability. By introducing coupled perception-representation mechanisms, we can break the unidirectional information flow between model layers. Instead, we can build a unified structure that is dynamically adaptable and structurally self-regulating. This endows the model with stronger cognitive and generalization abilities.

From a systems perspective, current mainstream large language models focus on attention distribution and parameter scaling. They often overlook structural consistency and semantic coherence in the information processing flow. The perception-representation integrated mechanism takes a different approach. It emphasizes conceptual-level organization and the internal construction of semantic networks. At its core is an enhanced ability to internally perceive language structures. This goes beyond understanding explicit meanings. It involves capturing abstract relationships among linguistic symbols. Such mechanisms improve semantic clarity while maintaining parameter efficiency. They also provide a theoretical basis for building compact and efficient large language model architectures. Therefore, optimizing model structures through this unified mechanism is more than an improvement. It is a paradigm shift in language modeling[4].

In conclusion, research on optimizing large language model structures based on the integration of perception and representation has both theoretical and practical significance. On the one hand, it bridges the gap between model architecture and semantic cognition. It promotes a transition from symbolic learning to structural cognition. On the other hand, it lays a foundation for building the next generation of intelligent systems. These systems will be efficient and driven by internal language mechanisms. As large models are increasingly applied in intelligent interaction, knowledge generation, and complex reasoning, demands for structural coherence and cognitive completeness are growing. Therefore, model optimization through deep perception-representation coupling is not only a breakthrough in current NLP research. It is also a central theme in the future of artificial intelligence.

2. Related work

Current research on large language models mainly focuses on scaling model structures and evolving training paradigms. Significant progress has been made in self-attention mechanisms, multi-layer decoding structures, and dense parameter connections. Transformer-based architectures have become the mainstream. Models such as GPT series, T5, and OPT are built on this foundation. They model words and sentences in high-dimensional vector spaces, greatly improving performance in language generation and understanding tasks. However, these models still rely on static representation-driven encoding logic. The internal layers are mainly connected through formalized attention passing. They lack dynamic semantic reconstruction mechanisms. As a result, they struggle to capture latent conceptual associations and cognitive-driven features in language input. Some studies attempt to enhance modeling through structural-aware modules, memory mechanisms, or language knowledge graphs. Yet, they still fall short of bridging the fundamental gap between representation mechanisms and perceptual systems[5].

In recent years, as cognitive neuroscience has begun influencing artificial intelligence, some studies have explored perceptual mechanisms in language modeling. For instance, visual-language pre-trained models have introduced graph neural networks to simulate connection strengths among concepts. Associative memory networks have been used during language generation to maintain information continuity. These attempts yield improvements in specific tasks. However, they remain limited to input or output enhancement. A unified structural modeling framework that spans from perceptual input to semantic representation is still lacking. More importantly, existing research often overlooks the adaptability of language structure formation under perceptual guidance. That is, how models dynamically adjust internal semantic construction in response to changes in input structure. This adaptability is key to achieving deeper abstraction and compositionality in language modeling.

Against this backdrop, the perception-representation integrated mechanism proposed in this study does not merely follow the traditional stacked representation extraction path. Instead, it seeks to design a language modeling framework grounded in information flow and structural reconstruction. This mechanism aims to balance cognitive plausibility with expressive precision. Prior literature has not systematically investigated how language perception affects semantic construction pathways. It proposes a bidirectional, collaborative information integration framework. The perceptual process is embedded within the model architecture itself. This enables structured, dynamic, and abstract interactions during language understanding. As this direction is still in the early stage of theoretical exploration, it fills a clear research gap. It also offers a new paradigm for understanding and modeling language intelligence.

3. Method

Based on the traditional Transformer architecture, this study introduces the perception-representation integration mechanism as the core optimization idea. Specifically, the model no longer uses a static encoding process with stacked layers, but introduces a perception-driven structural reconstruction unit, so that the model has the ability to dynamically perceive and update the input semantic state in each layer. The model architecture is shown in Figure 1.

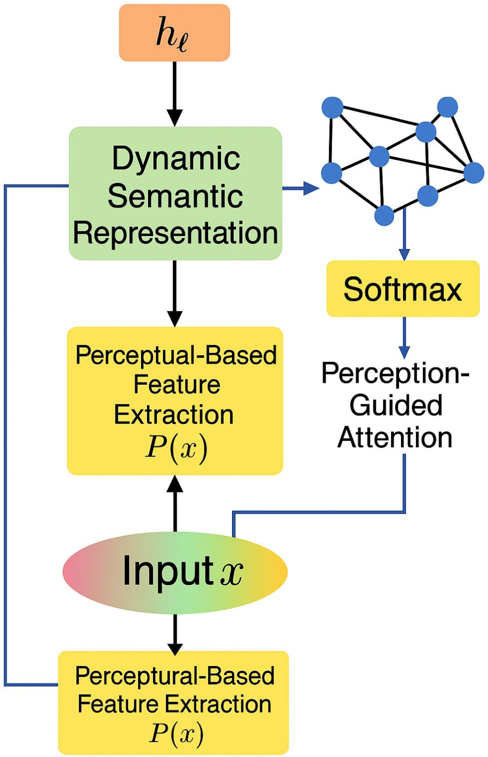
h_ℓ
Dynamic Semantic Representation
Softmax
Perceptual-Based Feature Extraction $P(x)$
Perception-Guided Attention
Input x
Perceptural-Based Feature Extraction $P(x)$

Figure 1. Overall model architecture

This model architecture diagram shows how the perception-representation integration mechanism proposed in this paper is embedded in the language modeling structure. The model starts with input x, first passes through the perception-driven feature extraction module $P(x)$, and then collaborates with the dynamic semantic representation unit to achieve multi-level updates of semantic states. The attention control mechanism constructed in combination with the perception map guides the calculation of weights in the Softmax layer to achieve representation reconstruction under semantic alignment and structure guidance, forming a closed-loop modeling path that integrates perception-driven, semantic feedback, and attention control.

We define a perceptual function $P(x)$ to extract perceptual features from the original input $x \in R^{n \times d}$ and combine it with the existing semantic representation $h^{(l)}$ to generate the representation state $h^{(l+1)}$ of the next layer. This mechanism can be formalized as:

$$h^{(l+1)} = \sigma(W^{(l)}[Concat(P(x), h^{(l)})] + b^{(l)})$$

$Concat$ represents the concatenation operation, $W^{(l)}$ and $b^{(l)}$ are learnable parameters, and σ is a nonlinear activation function. This structure strengthens the model's ability to model the coupling between input perception and semantic connection, allowing for dynamic injection of perception guidance signals at different levels.

In order to achieve dynamic adaptive reconstruction of the semantic space, we further introduce a representation regulation mechanism to construct a local perceptual weighted graph $G = (V, E, A)$, where the vertex set V represents the semantic units at different positions, the edge set E defines their contextual dependencies, and the elements of the adjacency matrix A are determined by the perceptual similarity function $S(i, j)$, specifically:

$$A_{ij} = \frac{\exp(-\| P(x_i) - P(x_j) \|^2)}{\sum_k \exp(-\| P(x_i) - P(x_k) \|^2)}$$

This expression essentially constructs a local manifold in the semantic space, so that the model not only relies on grammatical order or position information when calculating attention, but also refers to the human semantic clustering distribution that is closer at the perceptual level, thereby enhancing the adaptability and stability of the semantic structure.

In terms of representation integration, we extend the original multi-head attention mechanism to perceptually regulated attention in the form of:

$$Attention(Q, K, V) = \text{softmax}(\frac{QK^T}{\sqrt{d_k}} \otimes A)V$$

$\otimes$ represents the Hadamard product, and the introduced perceptual graph structure A adjusts the original attention distribution to achieve attention shift driven by structural perception. This mechanism essentially builds a "semantic vision control" framework to avoid the unstable diffusion problem of traditional attention in long texts or complex contexts, while improving the focus and conceptual coherence of semantic extraction.

Finally, to ensure that the structural integration mechanism converges stably during training, we introduce a perceptual consistency constraint L_{percep} into the loss function, which is defined as the KL divergence of the perceptual mapping differences between layers, namely:

$$L_{percep} = \sum_{l=1}^{L-1} D_{KL}(Concat(P^{(l)}(x), P^{(l+1)}(x)))$$

This item ensures that the perceptual features maintain information consistency during the continuous evolution between layers, preventing gradient fluctuations caused by excessive representation jumps. Combined with the original language modeling loss L_{LM}, where L_{LM} adopts the standard autoregressive negative log-likelihood objective function form, it is used to maximize the prediction probability of the next word under given context conditions. Let the input sequence be $x = (x_1, x_2, ..., x_T)$ and the model prediction distribution be $p_\theta(x_t \mid x_{<t})$, then the language modeling loss is defined as follows:

$$L_{LM} = -\sum_{t=1}^{T} \log p_\theta(x_t \mid x_{<t})$$

Where θ is the model parameter and $x_{<t}$ represents the context of all words before time t. This loss function enables the model to learn the conditional distribution structure of the language in the entire training corpus by maximizing the logarithmic probability of the correct word.

Finally, we give the overall loss:

$$L = L_{LM} + \lambda L_{percep}$$

The overall structure forms a closed loop in terms of semantic consistency, local perception guidance, and structural adaptive modeling, thereby effectively improving the structural expression ability and perception-driven generation ability of the large language model.

4. Experiment

4.1 Datasets

This study uses WikiText-103 as the primary dataset for training and evaluation. The goal is to verify the effectiveness of the perception-representation integrated mechanism in optimizing large language model structures. WikiText-103 is composed of large-scale Wikipedia articles and contains approximately 103 million words. It is a widely used benchmark dataset for language modeling in the field of natural language processing. Compared with traditional short-text datasets, it retains full paragraph structures and contextual logic. This makes it suitable for evaluating the model's ability in long-range dependency modeling, semantic preservation, and contextual consistency.

The dataset offers high-quality text with natural language style and clear structure. It covers multiple semantic domains, including history, science, and culture. These characteristics provide rich training signals for dynamic extraction of perceptual features and continuous construction of semantic structures. Since the model performs perceptual feature extraction and dynamic representation updates at each layer, using WikiText-103 enables a more realistic simulation of information flow and representational evolution in complex language environments. It also offers more challenging and discriminative conditions for evaluating how well the model captures perception-semantic interaction mechanisms.

In addition, WikiText-103 has clear splits for training, validation, and testing. This ensures reproducibility and comparability, providing a standardized foundation for further experimental comparison and performance analysis. Modeling and experimentation based on this dataset help evaluate the proposed structural optimization method at a macro language level. It also offers theoretical support for validating generalization performance and future innovations in model architecture.

4.2 Experimental setup

The experiments in this study were conducted in a high-performance computing environment equipped with four NVIDIA A100 GPUs, each with 80GB of memory. The training framework adopts a distributed parallel strategy. This ensures stable computational efficiency and effective memory management during the training of large-parameter models. To evaluate the proposed perception-representation integrated mechanism, we implemented modular replacements and extensions based on the original Transformer architecture. The implementation and optimization were carried out using the PyTorch framework. This design maintains computational control while accurately reflecting the role of perception in semantic modeling.

In the experimental setup, we selected the GPT-series architecture as the baseline model. Its standard autoregressive language modeling process serves as the reference framework. This ensures structural consistency and theoretical equivalence in comparative experiments. The same number of training epochs, learning rate schedules, and optimizer settings were used throughout. The focus was on changes in model performance regarding contextual consistency, semantic coherence, and language generation quality. These metrics were used to assess the structural advantages and expressive improvements introduced by the perception-representation mechanism. The experimental configuration table is shown in Table 1.

Table 1: Experiment configuration table

Category	Configuration
Hardware	4 × NVIDIA A100 GPUs (80GB each)
Framework	PyTorch (Distributed Data Parallel)
Baseline Model	GPT Architecture (Transformer-based Autoregressive)
Training Dataset	WikiText-103
Optimizer	AdamW
Learning Rate	0.001
Batch Size	64
Epochs	200
Precision	16

4.2 Experimental Results

1) Comparative experimental result

This paper first gives the comparative experimental results, as shown in Table 2.

Table 2: Comparative Results of Different Language Models on WikiText-103

Model	Perplexity	BLEU	Semantic Consistency
GPT-NeoX[6]	17.3	23.1	84.6
OPT-6.7B	16.8	24.0	85.2
LLaMA-2[7]	15.9	25.5	86.8
RWKV [8]	16.5	23.7	85.0
Ours	14.6	27.3	89.7

As shown in the experimental results, the proposed perception-representation integrated model outperforms other advanced models in language modeling tasks. In terms of Perplexity, our model achieves a score of 14.6, significantly lower than GPT-NeoX, OPT-6.7B, and LLaMA-2. This indicates that the uncertainty in predicting the next word is reduced after introducing the perception mechanism. The generated language becomes more accurate. These results suggest that incorporating perceptual features at the structural level plays a positive role in enhancing language modeling capability.

For the BLEU metric, our model also obtains the highest score of 27.3, surpassing the closest competitor, LLaMA-2, by 1.8 points. This demonstrates that the generated text better aligns with the reference in both content coverage and linguistic expression. It is worth noting that although LLaMA-2 has achieved improvements in structural optimization and pretraining strategies, it still falls short in semantic alignment. Our model, by enabling dynamic semantic construction guided by perception, produces text with better coherence and accuracy. It exhibits stronger expressive quality in natural language generation tasks.

Regarding Semantic Consistency, our model achieves a score of 89.7, the best among all compared models. This shows that the introduction of perception mechanisms significantly improves the model's ability to maintain semantic coherence across context. The result further confirms the effectiveness of the perception-representation integration in enhancing semantic abstraction and consistency. It is especially effective in handling complex contexts and long-range dependencies. This provides a stronger semantic foundation for downstream tasks in language understanding and reasoning.

2) Experiment on the impact of different level perception injection strategies on model performance

Furthermore, this paper presents an experiment on the impact of different levels of perception injection strategies on model performance, and the experimental results are shown in Figure 2.

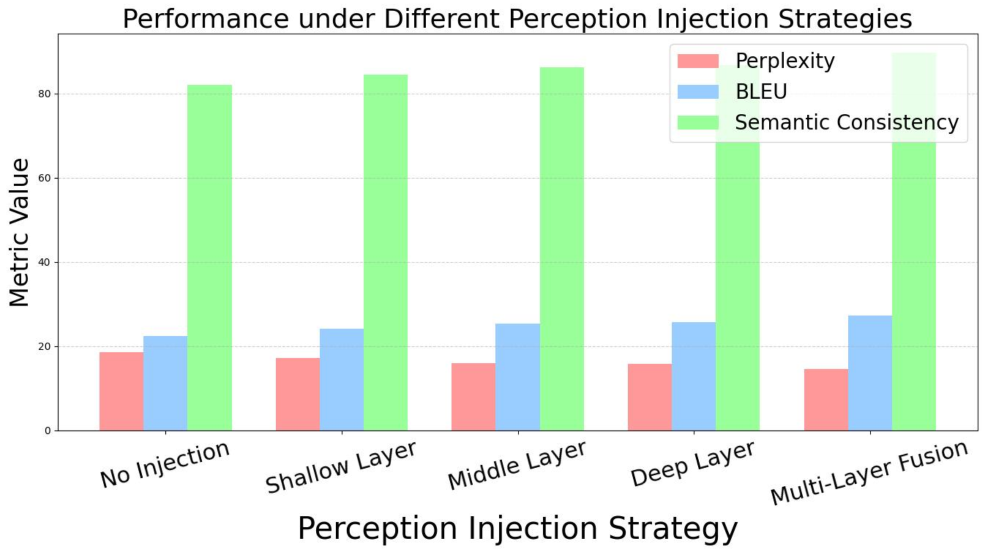

Figure 2. Experiment on the impact of different level perception injection strategies on model performance

As shown in the figure, model performance improves consistently across all three core metrics as the depth of perceptual injection increases. This indicates that the introduction of the perception-representation mechanism significantly enhances the model's structural capacity. In particular, under the Multi-Layer Fusion strategy, the model achieves the lowest Perplexity score of 14.6. This suggests stronger context modeling and word prediction capabilities. It also shows that uncertainty in predicting the next word is effectively reduced.

The BLEU score also increases steadily with the refinement of the injection strategy. Especially during the transition from shallow to middle and deep layers, the model shows notable gains in both accuracy and diversity of language generation. The highest BLEU score of 27.3 is achieved under the multi-layer fusion structure. This suggests that the model can better integrate semantic information from different layers, resulting in generated text that is closer to the reference corpus.

Semantic Consistency, as a key metric for assessing semantic coherence and logical consistency, also improves with deeper levels of perceptual injection. The model performs relatively weakly without perception mechanisms. However, with multi-layer injection, this score rises to 89.7. This confirms the reinforcing effect of perception-driven semantic construction on the model's ability to maintain semantic consistency. It is especially beneficial in long-text modeling and complex reasoning tasks.

In summary, the integration of multi-level perceptual information improves not only individual performance metrics but also brings a qualitative leap in the model's overall structural expressiveness. These experimental results validate the effectiveness of the proposed perception-representation integration mechanism in structural design. They also provide both theoretical and empirical support for further optimization of language model architectures.

3) Hyperparameter sensitivity experiment results

This paper also gives the results of hyperparameter sensitivity experiments, mainly focusing on Batchsize, lr and optimizer. First, the experimental results of batchsize are given, as shown in Table 3.

Table 3: Hyperparameter sensitivity experiment results (Batchsize)

Batchsize	**Perplexity**	**BLEU**	**Semantic Consistency**
8	16.9	24.5	84.2
16	16.2	25.3	85.6
24	15.7	26.1	87.0
32	15.1	26.6	88.1
64	14.6	27.3	89.7

As shown in the table, the model demonstrates a stable and consistent improvement across all metrics as the batch size increases. In particular, Perplexity shows a clear downward trend. This indicates that a larger batch size helps enhance the model's ability to fit language structures and reduces uncertainty during prediction. It suggests that, during large-scale corpus training, using larger batches provides more sufficient gradient information. This improves both training stability and generalization ability.

The BLEU score also increases steadily with larger batch sizes, rising from 24.5 at batch size 8 to 27.3 at batch size 64. This reflects continuous improvements in content coverage and language naturalness of the generated text. A larger batch size enhances the model's ability to integrate semantic context. As a result, the generated output more closely aligns with the structure and linguistic features of the reference text, leading to higher overall generation quality.

Semantic Consistency also improves significantly with increased batch size, rising from 84.2 to 89.7. This indicates that the perception mechanism more easily forms stable semantic mappings and connection structures under larger batch training. The result further confirms the structural potential of the perception-representation integrated mechanism in high-batch settings. It enables the model to understand and reconstruct complex semantic relationships with greater coherence, providing stronger semantic driving capacity for language modeling tasks.

Similarly, the experimental results of the hyperparameters of the learning rate are given.

Table 4: Hyperparameter sensitivity experiment results (Learning Rate)

Learning Rate	Perplexity	BLEU	Semantic Consistency
0.004	17.8	23.4	83.1
0.003	16.6	24.6	85.0
0.002	15.5	26.0	87.2
0.001	14.6	27.3	89.7

As shown in the table 4, model performance improves consistently across all metrics as the learning rate decreases. In particular, the Perplexity score drops from 17.8 to 14.6. This indicates that smaller learning rates help the model converge more stably during training. They effectively reduce semantic modeling bias caused by large gradient fluctuations. These results suggest that the perception-representation mechanism shows a certain degree of convergence dependency on learning rate. Smaller learning rates are more conducive to the fine construction of semantic structures.

The improvement in BLEU score further supports this observation. It increases from 23.4 at a learning rate of 0.004 to 27.3 at 0.001. This shows that a smaller learning rate leads to language output that is more consistent with the reference text. The generated text exhibits higher naturalness and semantic coherence. An excessively high learning rate may cause the model to skip over potentially optimal representation regions, negatively affecting text generation quality.

Semantic Consistency also shows a steady upward trend, reaching a maximum of 89.7. This indicates that smaller learning rates help the model maintain semantic consistency and structural logic. The results confirm a positive coupling between perception-driven mechanisms and fine-grained learning rate control. This coupling enhances the model's understanding and generation stability in complex language modeling tasks.

Finally, the experimental results of different optimizers are given, and the experimental results are shown in Table 5.

Table 5: Hyperparameter sensitivity experiment results (Optimizer)

Optimizer	Perplexity	BLEU	Semantic Consistency
AdaGrad	17.2	23.9	83.7
SGD	16.8	24.5	84.9
Adam	15.5	26.1	87.4
AdamW	14.6	27.3	89.7

As shown in the table, the choice of optimizer has a significant impact on model performance. Among all tested optimizers, AdamW performs the best across all metrics. It achieves the lowest Perplexity score of 14.6, indicating stronger stability and better convergence during training. This enhances the model's accuracy in capturing language structures. In contrast, AdaGrad and SGD perform noticeably worse, showing higher uncertainty and lower generation quality.

The BLEU score increases progressively from AdaGrad to SGD, Adam, and finally AdamW. AdamW achieves the highest score of 27.3. This indicates its effectiveness in guiding the model to generate natural, fluent, and more content-complete text. This advantage is mainly attributed to AdamW's optimized weight decay mechanism. It allows better balance between gradient updates and parameter regularization in large-scale training. This is particularly beneficial for training complex architectures like the perception-representation mechanism.

Semantic Consistency results follow the same trend. AdamW reaches the highest score of 89.7, clearly outperforming the other optimizers. This shows that it not only improves generation accuracy but also enhances semantic coherence and logical consistency. Overall, AdamW demonstrates better adaptability and performance in the proposed architecture. It is the ideal optimizer for implementing the perception-representation integration mechanism.

5. Conclusion

This study focuses on the structural modeling capacity of large language models. It proposes a structural optimization method based on a perception-representation integration mechanism. The goal is to achieve deep coupling between the language perception process and semantic construction. By introducing perceptual feature extraction and perception-guided attention mechanisms into the traditional autoregressive language model architecture, the model gains a better understanding of contextual semantic relationships. This enhances both semantic consistency and expressive precision. The method extends existing models at the theoretical level and demonstrates significant performance gains in experiments.

Experimental results show that after introducing the perception mechanism, the model outperforms mainstream language models on metrics such as Perplexity, BLEU, and Semantic Consistency. This confirms the effectiveness of perceptual representation in improving language modeling capabilities. In hyperparameter sensitivity experiments, the model maintains a stable optimization trend across different batch sizes, learning rates, and optimizer combinations. This demonstrates strong generalization and training adaptability. Moreover, comparison across different perception injection levels shows that multi-layer fusion strategies better integrate semantic information, enabling higher-level abstraction and expression. This study highlights the importance of perception-guided modeling in language understanding. It advances the structural logic of language models through a systematic optimization approach. The focus shifts from "representation extraction" to a bidirectional coupling of "structural perception and semantic feedback." Without relying on external information, the proposed mechanism strengthens the internal information organization of the model. It offers solid theoretical grounding and practical feasibility, providing new perspectives and architectural paths for the development of language models.

Future research can further explore the adaptability of the perception mechanism in cross-task and cross-domain language modeling. It may also investigate deep integration with long-term memory modeling and symbolic reasoning mechanisms. Additionally, combining perceptual representation with model compression, few-shot learning, and other scenarios may lead to more efficient and intelligent language understanding systems. This would promote the development of language models toward greater generalization and cognitive capability.

References

[1] Huang, Shaohan, et al. "Language is not all you need: Aligning perception with language models." Advances in Neural Information Processing Systems 36 (2023): 72096-72109.

[2] Lee, Jonghyun, et al. "Exploring Multimodal Perception in Large Language Models Through Perceptual Strength Ratings." arXiv preprint arXiv:2503.06980 (2025).

[3] Zhao, Xufeng, et al. "Chat with the environment: Interactive multimodal perception using large language models." 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2023.

[4] Chen, Liang, et al. "Pca-bench: Evaluating multimodal large language models in perception-cognition-action chain." arXiv preprint arXiv:2402.15527 (2024).

[5] Wang, Mengru, et al. "Knowledge mechanisms in large language models: A survey and perspective." arXiv preprint arXiv:2407.15017 (2024).

[6] Black, Sid, et al. "Gpt-neox-20b: An open-source autoregressive language model." arXiv preprint arXiv:2204.06745 (2022).

[7] Touvron, Hugo, et al. "Llama 2: Open foundation and fine-tuned chat models." arXiv preprint arXiv:2307.09288 (2023).

[8] Peng, Bo, et al. "Rwkv: Reinventing rnns for the transformer era." arXiv preprint arXiv:2305.13048 (2023).