Team narrators at PAN 2026: StyleGate: Dual-Branch LoRA on DeBERTa-v3 for Multi-Author Writing Style Analysis

PAN at CLEF 2026

Abdallah Elsayed[1,*], Elbadry Mohammed[1], Nagwa El-Makky[1] and Marwan Torki[1]

[1]*Alexandria University, Alexandria, Egypt*

Abstract

We present a dual-branch gated LoRA system for the PAN 2026 Multi-Author Writing Style Analysis (MAWSA) task, which requires detecting author-change boundaries in documents written by multiple authors across three difficulty levels. Our approach builds on DeBERTa-v3-large as a shared cross-encoder backbone and attaches two independent Low-Rank Adaptation (LoRA) adapters: a topic branch trained on easy-difficulty data, where topical shifts dominate author-change signal, and a style branch trained on medium- and hard-difficulty data, where fine-grained stylometric cues are essential. A lightweight MLP gate learns to dynamically blend the two branch predictions per sentence pair. Training proceeds in three sequential phases: (1) topic branch training on easy data, (2) style branch training on medium and hard data with the topic branch frozen, and (3) gate training on all difficulties with both branches frozen. On the official PAN 2026 test evaluation, our method achieves F1-macro scores of 0.968, 0.560, and 0.447 on the easy, medium, and hard sub-tasks, respectively.

Keywords

style change detection, authorship analysis, DeBERTa, LoRA, mixture of experts, multi-author writing,

1. Introduction

Identifying the points in a document where the author changes is a fundamental problem in computational authorship analysis [1]. The PAN 2026 Multi-Author Writing Style Analysis (MAWSA) task, part of the PAN 2026 lab [2], formalises this as a binary classification problem over consecutive sentence pairs (s_i, s_{i+1}): given a multi-author document, the goal is to predict whether the authorship changes between adjacent sentences. The three difficulty levels (easy, medium, and hard) differ in how strongly topic shifts and writing-style patterns indicate author boundaries.

A major challenge is that the three difficulty levels have different rates of author changes (30.5% for easy, 4.4% for medium, and 20.9% for hard) and depend on different types of signals. Easy documents often contain clear topic changes at author boundaries, while medium and hard documents rely more on writing-style cues. Consequently, a single fine-tuned model must generalise across different author-change patterns in all three sub-tasks.

We address this challenge using a parameter-efficient dual-branch architecture. Two LoRA adapters [3] share the same frozen DeBERTa-v3-large [4] backbone, but each is trained on a different subset of the data. This allows one branch to focus more on topic-related patterns, while the other focuses on writing-style features. An MLP gate then learns how to combine the predictions of the two branches for each input sample. This design helps the model handle both topic-driven and style-driven author changes while keeping the number of trainable parameters relatively small.

Our design choices are motivated by recent shared-task experience. In the previous edition of the task, PAN 2025 [5], DeBERTa-based cross-encoders were among the strongest approaches for multi-author writing style analysis; we therefore build on DeBERTa and adopt its most recent variant, DeBERTa-v3-large [4], which improves on earlier versions through ELECTRA-style pre-training and gradient-disentangled embedding sharing. Rather than fully fine-tuning this large backbone, we use LoRA [3] because it keeps the backbone frozen and trains only a small number of low-rank matrices,

CLEF 2026 Working Notes, 21 – 24 September 2026, Jena, Germany

*Corresponding author.

abdallah.elsaid250@gmail.com (A. Elsayed)

Table 1
Dataset statistics per difficulty level.

Difficulty	Train pairs	Val pairs	Positive %
Easy	549,804	117,435	30.5
Medium	1,299,324	278,988	4.4
Hard	589,946	127,380	20.9

which both reduces the risk of overfitting on the imbalanced sub-tasks and lets us maintain two specialised branches at a fraction of the cost of two full models. The per-branch ranks ($r = 16$ for the topic branch, $r = 32$ for the style branch) were chosen so that the style branch, which must capture more subtle stylometric cues on the medium and hard data, is given greater expressive capacity than the topic branch, whose easy-difficulty signal is largely topical and thus lower-rank. We do not explore alternative backbones such as large language models (LLMs) or other encoders in this work; we leave such comparisons to future experiments.

The remainder of this paper is structured as follows. Section 2 describes the task and dataset. Section 3 presents our system architecture. Section 4 details the three-phase training procedure. Section 5 reports experimental results, and Section 6 concludes.

2. Task and Dataset

2.1. Task Definition

The MAWSA task [6] provides a set of documents, each written by multiple authors. For a document containing n sentences, the system must produce a binary label for every consecutive pair (s_i, s_{i+1}), $i = 1, \ldots, n-1$, where label 1 indicates an author change between s_i and s_{i+1}, and label 0 indicates the same author continues. Three sub-tasks are defined according to difficulty:

- **Easy:** documents with clear topical transitions at author boundaries; high positive rate (30.5%).
- **Medium:** longer documents with style variation; very low positive rate (4.4%).
- **Hard:** moderate document length with stylistically similar authors; intermediate positive rate (20.9%).

2.2. Dataset Statistics

Table 1 summarises the corpus. Each difficulty level provides 10,500 training documents and 2,250 validation documents. Sentence pairs are formed from all adjacent sentences within each document and stored in a LMDB binary cache for fast loading during training. DeBERTa tokenisation is applied on the fly during batch collation with a maximum sequence length of 256 tokens.

3. System Architecture

3.1. Base Encoder

We use DeBERTa-v3-large [4] as a cross-encoder, motivated by the strong performance of DeBERTa-based models on the previous PAN 2025 edition of the task [5]. A cross-encoder is a natural fit for this pairwise setting because it lets the two sentences attend to each other directly, which is important for detecting the subtle style differences that separate same-author from different-author pairs. Each sentence pair (s_i, s_{i+1}) is formatted as a single sequence:

$$[\text{CLS}]\ s_i\ [\text{SEP}]\ s_{i+1}\ [\text{SEP}]$$

and fed to the encoder. The 1024-dimensional hidden state of the [CLS] token is used as the sequence representation. The base DeBERTa weights are kept frozen throughout all three training phases; only the LoRA adapters and downstream heads are trained.

3.2. Dual LoRA Adapters

LoRA [3] injects trainable low-rank matrices into existing weight matrices without modifying the original parameters. For a weight matrix $W \in \mathbb{R}^{d \times k}$, LoRA adds:

$$W' = W + BA, \quad B \in \mathbb{R}^{d \times r},\ A \in \mathbb{R}^{r \times k} \tag{1}$$

where $r \ll \min(d, k)$ is the rank. We apply LoRA to the query, key, and value projection layers of all attention heads. We prefer LoRA over full fine-tuning because it freezes the backbone and trains only a small set of low-rank matrices, which reduces overfitting on the imbalanced sub-tasks and makes it inexpensive to maintain two specialised branches on a single shared encoder.

We maintain two separate LoRA adapters. The ranks are set so that the style branch, which must model finer-grained stylometric cues, has more capacity than the topic branch, whose easy-difficulty signal is dominated by lower-rank topical structure:

- **Topic branch** (rank $r = 16$, scaling $\alpha = 32$): trained on easy data. This branch develops representations sensitive to content and topic shifts.
- **Style branch** (rank $r = 32$, scaling $\alpha = 64$): trained on medium and hard data. A higher rank is used to allow greater expressive capacity for fine-grained stylometric features.

Each branch has its own combiner network and classification head:

$$h_{\text{branch}} = \text{GELU}\big(\text{Linear}_{1024 \to 512}([\texttt{CLS}])\big), \quad \ell_{\text{branch}} = \text{Linear}_{512 \to 1}(h_{\text{branch}})$$

3.3. MLP Gate

The gate module receives the hidden representations from both branches and produces a scalar mixture weight $\alpha \in (0, 1)$:

$$\alpha = \sigma\Big(\text{Linear}_{256 \to 1}\big(\text{ReLU}\big(\text{Linear}_{1024 \to 256}([h_{\text{topic}}; h_{\text{style}}])\big)\big)\Big) \tag{2}$$

where $[\cdot\,;\cdot]$ denotes concatenation. The final prediction logit is:

$$\ell = \alpha \cdot \ell_{\text{topic}} + (1 - \alpha) \cdot \ell_{\text{style}} \tag{3}$$

When $\alpha \to 1$ the model relies on the topic branch; when $\alpha \to 0$ it relies on the style branch. The gate therefore acts as a learned, per-sample difficulty estimator: it routes easy pairs toward the topic branch and challenging pairs toward the style branch without any explicit supervision on which branch to trust.

4. Three-Phase Training

Training is divided into three sequential phases, each with a clearly defined scope of trainable parameters. This progressive strategy ensures that each component is initialised from a well-trained predecessor before the next component is added.

Table 2
Shared training hyperparameters.

Hyperparameter	Value
Optimiser	AdamW
LoRA / encoder LR	1×10^{-4}
Head / gate LR	2×10^{-4}
Weight decay	0.01
Scheduler	CosineAnnealingLR
$\eta_{\min}$	1×10^{-7}
Gradient clip norm	1.0
Mixed precision	bfloat16
Max sequence length	256 tokens

Table 3
F1-macro scores on the TIRA evaluation platform and internal F1-micro scores for the Phase 1 topic branch.

Sub-task	F1-macro (TIRA)	F1-micro (internal)
Task 1 (Easy)	0.968	1.000
Task 2 (Medium)	0.560	–
Task 3 (Hard)	0.447	–

4.1. Phase 1 — Topic Branch

The topic LoRA adapter, its combiner, and its classification head are trained on the easy sub-task only. We use weighted binary cross-entropy with pos_weight = 2.3 to compensate for the class imbalance. Training uses the AdamW optimiser with a cosine annealing schedule, and early stopping is applied when validation F1-macro ceases to improve (patience = 5 epochs). The topic branch achieves a perfect F1-micro of 1.000 on both the training and validation splits of the easy sub-task, confirming that topical author-change signals are fully captured by this branch. After Phase 1 the topic branch weights are frozen.

4.2. Phase 2 — Style Branch

The topic branch checkpoint is loaded. The base encoder and topic branch remain frozen. A fresh style LoRA adapter, combiner, and classification head are initialised and trained on the combined medium and hard data. We use pos_weight = 2.0 and train for up to 20 epochs with early stopping (patience = 5, monitor: val/F1-macro). The style branch is frozen after Phase 2.

4.3. Phase 3 — Gate

The Phase 2 checkpoint is loaded. Both LoRA branches, their combiners, and their classification heads are frozen. Only the MLP gate is trained, using all three difficulty levels combined with pos_weight = 4.4 and up to 10 epochs (patience = 5). Training all three difficulties jointly exposes the gate to the full range of easy-to-hard samples so it can calibrate α appropriately for each regime.

4.4. Training Hyperparameters

Table 2 summarises the key hyperparameters shared across phases.

5. Results

Table 3 reports F1-macro scores on the TIRA evaluation platform [7] for each sub-task.

The topic branch (Phase 1) achieves a perfect F1-micro of 1.000 on both the training and validation splits of the easy sub-task, demonstrating that DeBERTa with a lightweight LoRA adapter can fully exploit the strong topical signals present in easy documents. The final gated system carries this strength through to evaluation, yielding a near-perfect TIRA F1-macro of 0.968 on Task 1. Performance drops substantially on the medium sub-task, reflecting the inherent difficulty of detecting author changes at a 4.4% positive rate where the class imbalance is severe. The hard sub-task, with its stylistically similar authors, poses a complementary challenge; the moderate F1-macro of 0.447 indicates that the style branch captures some discriminative signal but that this remains an open problem.

6. Conclusion

We have presented a dual-branch gated LoRA system for multi-author style change detection. By training two LoRA adapters on complementary data subsets and combining them with a lightweight MLP gate, the system can adapt its prediction strategy to the topical or stylometric nature of each sentence pair without modifying the shared DeBERTa backbone. The three-phase training procedure is simple to implement and keeps the total number of trainable parameters well below that of a full fine-tuned model.

Future work includes increasing the rank of the style branch LoRA adapter, exploring contrastive pre-training objectives to sharpen stylometric representations, and calibrating the decision threshold per difficulty level to better handle the severe class imbalance in the medium sub-task. Because the present system is built solely on a DeBERTa-v3 backbone, we also plan to compare it against alternative encoders and large language model (LLM) based approaches, which we did not explore in this work.

Declaration on Generative AI

During the preparation of this work, the author(s) used AI-based writing and coding assistance tools, namely Claude Sonnet 4.6 and ChatGPT, to support code development and manuscript drafting. All generated content was subsequently reviewed, edited, and verified by the author(s), who take full responsibility for the final publication.

References

[1] E. Stamatatos, A survey of modern authorship attribution methods, Journal of the American Society for information Science and Technology 60 (2009) 538–556.

[2] J. Bevendorff, M. Fröbe, A. Greiner-Petter, A. Jakoby, M. Mayerl, P. Nakov, H. Plutz, M. Potthast, B. Stein, M. N. Ta, Y. Wang, E. Zangerle, in: M. Hagen, M. Potthast, B. Stein, P. Schaer, E. Zangerle, A. Barrón-Cedeño, A. G. S. de Herrera, E. S. Salido, S. MacAvaney, J. M. Struß (Eds.), Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings of the Seventeenth International Conference of the CLEF Association (CLEF 2026), Lecture Notes in Computer Science, Springer, Berlin Heidelberg New York, 2026.

[3] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, W. Chen, et al., Lora: Low-rank adaptation of large language models., Iclr 1 (2022) 3.

[4] P. He, J. Gao, W. Chen, Debertav3: Improving deberta using electra-style pre-training with gradient-disentangled embedding sharing, arXiv preprint arXiv:2111.09543 (2021).

[5] J. Bevendorff, D. Dementieva, M. Fröbe, B. Gipp, A. Greiner-Petter, J. Karlgren, M. Mayerl, P. Nakov, A. Panchenko, M. Potthast, A. Shelmanov, E. Stamatatos, B. Stein, Y. Wang, M. Wiegmann, E. Zangerle, Overview of pan 2025: Generative ai detection, multilingual text detoxification, multi-author writing style analysis, and generative plagiarism detection: Extended abstract, in: Advances in Information Retrieval: 47th European Conference on Information Retrieval, ECIR 2025, Lucca, Italy, April 6–10, 2025, Proceedings, Part V, Springer-Verlag, Berlin, Heidelberg, 2025, p. 434–441. URL: https://doi.org/10.1007/978-3-031-88720-8_64. doi:10.1007/978-3-031-88720-8_64.

[6] E. Zangerle, M. Mayerl, M. Potthast, B. Stein, Overview of the Multi-Author Writing Style Analysis Task at PAN 2026, in: A. Barrón-Cedeño, A. G. S. de Herrera, E. S. Salido, S. MacAvaney, J. M. Struß (Eds.), Working Notes of CLEF 2026 – Conference and Labs of the Evaluation Forum, CEUR Workshop Proceedings, CEUR-WS.org, 2026.

[7] M. Fröbe, M. Wiegmann, N. Kolyada, B. Grahm, T. Elstner, F. Loebe, M. Hagen, B. Stein, M. Potthast, Continuous Integration for Reproducible Shared Tasks with TIRA.io, in: J. Kamps, L. Goeuriot, F. Crestani, M. Maistro, H. Joho, B. Davis, C. Gurrin, U. Kruschwitz, A. Caputo (Eds.), Advances in Information Retrieval. 45th European Conference on IR Research (ECIR 2023), Lecture Notes in Computer Science, Springer, Berlin Heidelberg New York, 2023, pp. 236–241. URL: https://link.springer.com/chapter/10.1007/978-3-031-28241-6_20. doi:10.1007/978-3-031-28241-6_20.