

Article

Not peer-reviewed version

---

# Redefining Textual Dynamics for Enhanced Text Style Transfer

---

Carlos Asanka , Conti Vatsalan , [Rodolfo Patel](#) \*

Posted Date: 4 December 2023

doi: 10.20944/preprints202312.0144.v1

Keywords: <span>Syntax-Aware Text Transformation;</span><span>&nbsp;</span><span>Style Classifier Analysis</span>



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: 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.

Article

# Redefining Textual Dynamics for Enhanced Text Style Transfer

Carlos Asanka, Conti Vatsalan and Rodolfo Patel \*

Briar Cliff University

\* Correspondence: rodolfopatel5@gmail.com

**Abstract:** Conventional text style transfer (TST) methodologies primarily utilize style classifiers to segregate the content and stylistic elements of text for effective style transformation. Despite the pivotal role of these classifiers, their influence on TST techniques remains largely unexplored. This study embarks on a detailed exploration of the limitations inherent in style classifiers within current TST frameworks. We reveal that these classifiers often inadequately comprehend sentence syntax, leading to diminished performance in TST models. In response, we introduce the Syntax-Enhanced Style Transfer (SEST) model, a groundbreaking approach incorporating a syntax-sensitive style classifier. This classifier ensures that the extracted style representations robustly encapsulate syntax nuances, enhancing TST effectiveness. Rigorous evaluations across diverse TST benchmarks demonstrate that SEST significantly surpasses contemporary models in performance. Additionally, our case studies highlight SEST's proficiency in producing syntactically coherent sentences that aptly retain original content.

**Keywords:** syntax-aware text transformation; style classifier analysis

## 1. Introduction

The field of Text Style Transfer (TST) has emerged as a significant area within the domain of natural language generation. Its primary focus is the alteration of stylistic aspects of text, such as tone or sentiment, while maintaining the original content's neutral style [1,2]. A unique challenge in TST is the requirement for training datasets that have the same content but exhibit varied stylistic elements, leading to a preference for unsupervised learning methods in most TST research. These approaches often utilize datasets that are not directly parallel but are annotated with stylistic markers.

One common methodology in TST research employs an adversarial learning framework with an autoencoder architecture. In this framework, a style classifier or discriminator is employed to initially differentiate between the content and the style elements of the text. Following this differentiation, a decoder is then used to reconstruct the text in a specified style [3–8]. Additionally, there are methods focused on attribute-controlled generation. These methods create a specific style attribute vector, which is then integrated with the latent representation of the text to produce outputs in a targeted style [9–13]. Similar to the adversarial learning model, the development of the style attribute vector in these methods is also guided by a pre-trained style classifier.

Both of these approaches in TST place significant emphasis on the style classifier. However, the depth of impact and the effectiveness of these classifiers in truly grasping the nuances of textual style, especially syntax [14], have not been thoroughly investigated. This paper seeks to fill this gap by presenting a comprehensive empirical examination of the role and efficacy of style classifiers in TST models.

Advancing from this detailed analysis, we propose the Syntax-Enhanced Style Transfer (SEST) model. This innovative model integrates a syntax-focused style classifier, ensuring a more nuanced incorporation of syntactic elements into the style representations for effective TST. Through extensive testing across a variety of TST datasets and augmented by human evaluation studies, we have found that SEST significantly outperforms existing top-tier models in this field. The principal contributions of this paper can be summarized as follows:

- A thorough empirical investigation into the capabilities of style classifiers in contemporary TST models, with a focus on their proficiency in understanding and integrating syntax.
- The development and introduction of the SEST model, a pioneering approach that places a heightened emphasis on the interplay between sentence structure and style representation learning.
- A series of rigorous experiments conducted on benchmark TST datasets, which clearly demonstrate the advanced performance and efficacy of SEST compared to other leading methods in the field.

## 2. Related Work

The exploration of text style within the realms of linguistic and computational studies has gained significant traction in recent years. The endeavor of Text Style Transfer (TST) is particularly captivating, as it seeks to modify the stylistic aspects of text, such as tone or diction, while ensuring the core, style-neutral content remains intact. A comprehensive and recent survey by Hu et al. [1] provides an extensive analysis of various techniques and progressions in the TST domain. In the early phases of TST research, the use of parallel corpora was the norm [18–23]. However, the scarcity of such datasets, especially in applications demanding diverse stylistic dialogue generation, has catalyzed the development of novel TST methodologies that circumvent the need for parallel data.

Within these novel methodologies, the extraction and manipulation of latent sentence representations stand out as a key strategy. Two dominant methods in this regard are: (1) adversarial learning and (2) attribute-controlled generation. Shen et al. [3] pioneered the adversarial learning method in TST, where the primary goal is for a classifier to evaluate an encoder's ability to generate content representations that are stylistically neutral. These representations are then fed into a style-specific decoder to produce text in the desired stylistic form. This adversarial approach has been further refined and diversified in subsequent research [4,6,8,24–30].

The attribute-controlled generation method, first proposed by Hu et al. [9], utilizes a Variational Autoencoder (VAE) [31] to learn a sentence's latent representation, denoted as  $z$ . This approach also incorporates a style classifier to extract a style attribute vector  $s$ , which, in conjunction with  $z$ , is used to generate text in the target style. This attribute-controlled approach has been echoed and adapted in several other TST studies [11,12,32]. The probabilistic encoder in the VAE serves an implicit role in differentiating style and content, ensuring that the manipulation of attribute codes does not result in a conflation of these two elements. This aspect of attribute control in TST has been a focal point in additional research endeavors [11,12,32,33].

The role of pretrained style classifiers in these methodologies cannot be overstated, as they are instrumental in steering the TST process. However, a critical observation is that these classifiers often overlook the syntactic aspects of sentences. Considering the profound impact of syntax on text style, especially in contexts requiring formal style transfer, this paper posits that incorporating syntactic considerations is paramount in TST. In this vein, we introduce the Syntax-Enhanced Style Transfer (SEST) method. SEST represents a breakthrough in TST research, as it not only acknowledges but integrates syntactic elements, thus achieving enhanced performance over current leading TST methods.

## 3. Preliminary Study

Prior to introducing our Syntax-Enhanced Style Transfer (SEST) method, we embarked on a detailed empirical study to scrutinize the capability of existing style classifiers in TST models in learning and distinguishing syntactic styles in text. Notable style classifiers such as TextCNN [34], RNN [35], and Transformer [36] have been widely utilized in various TST frameworks [11,30,32,37,38]. In our study, these classifiers were trained on the GYAFC dataset [39], a prominent dataset for formality transfer research. Initially, the classifiers were trained and tested with the standard GYAFC training and test sets. Subsequently, we introduced a structural variation to the GYAFC test set by randomly

altering the word order in sentences, theorizing that significant syntactic differences exist between formal and informal text, and such disruption should adversely affect classification accuracy.

The results, as shown in Table 1, underscore the pivotal role of syntax in determining text formality. A marginal decline of 2.9% in classification accuracy was observed in the *Scrambled* test set, indicating the influence of syntactic disruption. A deeper analysis revealed a marked decrease in accuracy for formal sentences upon alteration of their structure, contrasting with the relatively stable performance for informal sentences under similar conditions. This suggests that the classifiers might be prioritizing attribute words in their style determination, potentially overlooking crucial syntactic elements. Moreover, the classifiers appeared to categorize the syntactically altered sentences as informal, a misconception given the distinct syntax of genuinely informal sentences. This observation raises concerns about the classifiers' efficacy in discerning varied syntactic patterns associated with different formality styles. Crucially, such a limitation in understanding syntax could lead to the generation of incoherent sentences by the TST models, especially when adapting content to an informal style.

**Table 1.** Performance of style classifiers on the *GYAFC* test set and the modified *Scrambled* test set. **ACC** denotes overall accuracy for both formal and informal sentences, **F** for formal sentence accuracy, and **I** for informal sentence accuracy.

Classifier	Test set	ACC	F	I
TextCNN	GYAFC	88.6	91.3	86.4
	Scrambled	85.3	84.9	85.5
RNN	GYAFC	85.6	84.6	86.4
	Scrambled	82.2	74.8	87.8
Transformer	GYAFC	84.9	86.7	83.7
	Scrambled	82.9	80.5	84.6

#### 4. The Proposed Method

This section introduces the Syntax-Enhanced Style Transfer (SEST) model, crafted to overcome the limitations of existing TST methods in capturing and manipulating sentence structures during style transfer. We begin by discussing Graph Convolutional Networks (GCNs), followed by an explanation of how GCNs are employed in our syntax-focused classifier and encoder, the two critical components of the SEST model. Finally, we delineate the learning algorithm of SEST.

##### 4.1. GCN and Sentence Structure Representation

Graph Convolutional Networks (GCNs), a convolutional neural network variant [40] tailored for graph-structured data [41], have shown their prowess in leveraging syntactic dependency graphs for text representation [42]. For a graph  $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$  with nodes  $\mathcal{V}$  and edges  $\mathcal{E}$ , and a feature matrix  $X \in \mathbb{R}^{n \times d}$ , the GCN propagation rule is:

$$H^{(l+1)} = \sigma(AH^{(l)}W^{(l)}), \quad (1)$$

where  $H^{(l)}$  is the  $l$ -th layer feature matrix,  $W^{(l)}$  the weight matrix,  $A$  the adjacency matrix, and  $\sigma(\cdot)$  a non-linear activation function. GCNs input  $X$  and yield a latent feature matrix  $H^{(L)}$ , where  $L$  is the GCN layer count.

SEST aims to harness sentence structure information, critical in generating stylistically accurate sentences. Dependency trees, which represent syntactic relationships between words, can be graphically modeled and analyzed using GCNs [42,43]. To address over-parametrization in large datasets, we adopt a simplified adjacency matrix approach. This matrix encodes the dependency relations in a sentence, where columns and rows represent head words and dependents, respectively. Elements  $A_{ij}$  are set to 1 if a dependency exists. Self-loops are included for each node, following [42].

#### 4.2. Syntax-Aware Style Classifier

Our syntax-aware style classifier  $D$  is designed to better encode syntactic information from dependency trees. Sentence tokens of length  $n$ , denoted as  $s = \{w_1, \dots, w_n\}$ , are initially encoded in the word embedding layer. Considering GCN's limitation in capturing long-range dependencies, we apply graph convolution operations on Bi-LSTM hidden states rather than static embeddings [42]. These Bi-LSTM states,  $H_{lstm}$ , form the GCN input, where each  $h_{lstm,i}$  is a forward and backward hidden state concatenation. We apply a  $L$ -layer GCN, ensuring hidden representations are influenced by their neighbors within  $L$  edges in the dependency tree. The hidden representation at layer  $(l + 1)$  is given by:

$$h_i^{(l+1)} = \sigma\left(\sum_{j=1}^n A_{ij} W^{(l)} h_j^{(l)} + b^{(l)}\right) \quad (2)$$

Scaled dot-product attention [36] is then used to aggregate node representations into a cohesive sentence representation. The final style prediction employs a fully connected network and softmax operation on this aggregated representation.

#### 4.3. Syntax-aware Controllable Generation

The SEST model framework processes each input sentence  $s$  with attribute  $y_o$  and its corresponding adjacency matrix  $A$ . The syntax-aware encoder  $E$  encodes  $s$  into a latent representation  $z = E(s, A)$ .  $E$  extracts sentence structure using our classifier's feature extractor. The decoder  $G$  then reconstructs the sentence  $s = G(z, y_o)$  or generates a transferred sentence  $\tilde{s} = G(z, y_t)$ . Dependency trees for transferred sentences and their adjacency matrices  $\tilde{A}$  are generated using the Stanza parser [44]. The syntax-aware classifier  $D$  assesses the style of  $\tilde{s}$ . SEST is trained with a classification loss  $L_{cla}$  and a reconstruction loss  $L_{rec}$ .

**Classification Loss  $L_{cla}$ :** This loss ensures that the transferred sentence aligns with the target style. The pretrained syntax-aware classifier directs parameter updates for target-style prediction:

$$L_{cla} = -\mathbb{E}_{(s, y_o) \sim D} [\log P(y_t | G(\tilde{s}), \tilde{A})] \quad (3)$$

**Reconstruction Loss  $L_{rec}$ :** This loss preserves the original content in the transferred sentences. It is defined as:

$$L_{rec} = -\log P(s | z, y_o) \quad (4)$$

**Combined Loss:** The overall training loss  $L$  balances style transfer and content preservation:

$$L = L_{rec} + \lambda L_{cla} \quad (5)$$

Where  $\lambda$  is a hyper-parameter.

## 5. Experiments

### 5.1. Experimental Setup

This section presents the evaluation of the Syntax-Enhanced Style Transfer (SEST) method on two widely-recognized datasets. Our comprehensive experiments aim to benchmark SEST against 12 leading TST methods.

**Datasets.** The evaluation of SEST focuses on two critical style transfer tasks: (1) Sentiment transfer, and (2) Formality transfer. For sentiment transfer, we utilize the well-known Yelp restaurant review dataset [3], which comprises reviews classified as positive or negative based on their ratings. The GYAFC (Grammarly's Yahoo Answers Formality Corpus) dataset [39], specifically the *Family&Relationship* (F&R) domain, is employed for the formality transfer task. Table 2 delineates the dataset splits for both Yelp and GYAFC used in our experiments.

**Table 2.** Statistics of Yelp and GYAFC datasets for sentiment and formality style transfer tasks.

Dataset	Style	Training	Validation	Testing
Yelp	Positive	267K	38K	76K
	Negative	176K	25K	50K
GYAFC	Informal	51K	2.7K	1.3K
	Formal	51K	2.2K	1K

**Sentiment Transfer (Yelp).** The Yelp dataset, a compilation of restaurant reviews, serves as a prime resource for evaluating sentiment transfer, where the objective is to alter the sentiment of a sentence while retaining its contextual meaning. Following [3], reviews are classified based on a 5-point scale, with ratings above 3 labeled as positive and those below 3 as negative. Neutral reviews (rating of 3) are excluded.

**Formality Transfer (GYAFC).** The GYAFC dataset [39] is pivotal for assessing formality transfer, which involves transforming the tone of a sentence from informal to formal and vice versa. Formality transfer is intricate as it encompasses multiple text attributes like sentence structure, text length, punctuation, and capitalization. The dataset, consisting of manually rewritten informal sentences into their formal counterparts, provides a rich source for this task.

**Baselines.** SEST is benchmarked against a suite of 12 advanced TST models, including *ARAE* [4], *DualRL* [37], *DAST* and *DAST-C* [32], *PFST* [45], *DRLST* [30], *DeleteOnly*, *Template*, *Del&Retri* [46], *DIRR* [47].

**Training Configuration.** The experiments were conducted on a high-performance computing setup with Nvidia RTX 2080Ti GPUs. The word embeddings are 300-dimensional, learned from scratch. The SEST architecture comprises a single Bi-LSTM layer followed by 2 GCN layers, with the latent representation dimension set to 500. Style labels are encoded into 200-dimensional vectors. The decoder initializes by concatenating the latent representation  $z$  with the attribute controlling code  $y$ . A pre-trained syntax-aware style classifier assists in training, ensuring that the generated sentences align with the desired style. The Gumbel-softmax technique [48] is employed for back-propagation. The learning rate is set at  $1 \times 10^{-5}$ , and  $\lambda$ , the balance parameter, is set to 1.

## 5.2. Automatic Evaluation Metrics

The evaluation of SEST and baseline models is based on transfer strength, content preservation, and fluency.

**Transfer Strength.** The effectiveness of a TST model in achieving style modification is gauged through style transfer accuracy [1]. A pre-trained syntax-aware classifier determines the accuracy by predicting the style label of transferred sentences, considering the target style as the ground truth.

**Content Preservation.** Quantitative assessment of content retention post-transfer employs several metrics:

- *BLEU* and *self-BLEU* measure the similarity of transferred sentences with human references and their original versions, respectively.
- *Cosine Similarity* evaluates the semantic closeness between original and transferred sentences.
- *Word Overlap* quantifies the common unigram word rate between the original and transferred sentences.

**Fluency.** Fluent sentence generation is crucial for TST models. We utilize a fine-tuned GPT-2 model [49] to compute the perplexity (PPL) of transferred sentences, with lower PPL indicating higher fluency.

**G-Score.** The G-Score is a geometric mean of style transfer accuracy, BLEU, self-BLEU, cosine similarity, word overlap, and the inverse of perplexity, providing a comprehensive performance assessment.

### 5.3. Automatic Experiment Results

Table 3 presents the performance of the Syntax-Enhanced Style Transfer (SEST) model alongside other baseline models on the formality transfer task. SEST demonstrates superior G-Score performance, eclipsing other state-of-the-art methods. It's notable that most TST models show a trade-off between transfer strength and content preservation, but SEST manages to balance these aspects effectively, achieving 84.1% transfer accuracy and a BLEU score of 21.1. The GYAFC dataset also includes human reference performances, where SEST's metrics are closely aligned with these human benchmarks.

**Table 3.** Performance comparison on GYAFC dataset for formality transfer task.

Model	ACC(%)	BLEU	CS	WO	PPL	G-Score
ARAE [4]	76.2	2.2	0.903	0.042	35	0.71
DeleteOnly [46]	18.7	16.2	0.945	0.431	74	1.11
Template [46]	44.7	19.0	0.943	0.509	102	1.32
Del&Retri [46]	50.7	11.8	0.934	0.345	74	1.21
DualRL [37]	59.8	18.8	0.944	0.447	266	1.12
DAST [32]	78.3	14.3	0.934	0.350	352	1.01
DAST-C [32]	79.2	13.8	0.927	0.328	363	0.98
DRLST [30]	49.8	2.7	0.909	0.342	<b>31</b>	1.06
PFST [45]	48.3	16.5	0.940	0.393	116	1.25
DIRR [47]	71.8	18.2	0.942	0.451	145	1.28
ours	<b>84.1</b>	<b>21.1</b>	<b>0.962</b>	<b>0.591</b>	73	<b>1.69</b>
Human0	<b>84.6</b>	24.6	<b>0.942</b>	<b>0.393</b>	<b>24</b>	<b>2.00</b>
Human1	83.8	24.3	0.931	0.342	27	1.89
Human2	83.6	24.6	0.932	0.354	27	1.91
Human3	82.1	<b>24.7</b>	0.931	0.354	27	1.90

In the sentiment transfer task, similar trends are observed. Table 4 illustrates SEST's performance on the Yelp dataset, where it surpasses baselines in G-Score. The sentiment transfer task exhibits a higher average style transfer accuracy (86.3% for Yelp) compared to formality transfer, underscoring the complexity of the latter. Despite this, SEST effectively negotiates the trade-off between transfer strength and content preservation across both tasks.

**Table 4.** Performance comparison on Yelp dataset for sentiment transfer task.

Model	ACC(%)	self-BLEU	CS	WO	PPL	G-Score
ARAE [4]	83.2	18.0	0.874	0.270	79	1.35
DeleteOnly [46]	84.2	28.7	0.893	0.501	130	1.53
Template [46]	78.2	48.1	0.850	0.603	250	1.50
Del&Retri [46]	88.1	30	0.897	0.464	88	1.66
DualRL [37]	79.0	<b>58.3</b>	0.970	<b>0.801</b>	117	1.98
DAST [32]	90.7	49.7	0.961	0.705	181	1.76
DAST-C [32]	93.6	41.2	0.933	0.560	274	1.49
DRLST [30]	91.2	7.6	0.904	0.484	<b>65</b>	1.36
PFST [45]	85.3	41.7	0.902	0.527	94	1.78
DIRR [47]	<b>94.2</b>	52.6	0.957	0.715	292	1.63
SACG (ours)	93.0	57.7	<b>0.971</b>	0.778	74	<b>2.23</b>

### 5.4. Human Evaluation

A human evaluation study was conducted to assess the quality of sentences generated by SEST compared to leading baselines. A sample of 200 sentences from the GYAFC dataset underwent style transformation using SEST and four top-performing baselines. Two linguistics experts then evaluated these sentences based on transfer strength, content preservation, and fluency. The evaluators rated

content preservation and fluency on a 6-point Likert scale and identified whether the transformed sentences matched the target style.

Table 5 displays the results of this human evaluation. SEST excels in all three criteria, particularly in generating syntactically correct and fluent sentences. Inter-annotator agreement was calculated to minimize biases, with Cohen’s kappa coefficients indicating substantial agreement in content preservation and fluency, and moderate agreement in style transfer strength.

**Table 5.** Human evaluation scores on the GYAFC dataset.

Model	Style(%)	Content	Fluency
DualRL	28.5	4.09	4.52
DAST	27.5	3.22	3.68
PFST	24.0	3.91	4.54
Del&Retri	25.5	2.61	3.23
SEST	<b>44.5</b>	<b>4.39</b>	<b>5.07</b>

### 5.5. Syntax Evaluation

We conducted a syntax evaluation using the Tree Edit Distance (TED) on constituency trees, comparing TST model outputs with human references in the GYAFC dataset. This comparison aims to measure how closely the TST models align with human-like syntactic structures. Table 6 shows the results, where SEST outperforms baselines in producing sentences with syntactic structures similar to human references. This finding indicates SEST’s capability to grasp and replicate the syntactic nuances associated with different styles.

**Table 6.** Syntactic similarity (TED) between model outputs and human references in the GYAFC dataset.

Model	TED	Model	TED
DRLST	19.2	DeleteOnly	18.2
ARAE	18.1	Template	17.9
DualRL	15.2	Del&Retri	21.0
DAST	16.6	HPAY	18.4
PFST	15.5	DIRR	15.5
DAST-C	16.9	SEST	<b>13.2</b>

### 5.6. Ablation Study

An ablation study was conducted to evaluate the contribution of syntax-aware components in SEST. Table 7 outlines the study’s results, comparing the full SEST model against variants lacking the syntax-aware encoder and both the syntax-aware encoder and classifier. The results indicate that both components are crucial for maintaining high performance, particularly in terms of fluency and content preservation.

**Table 7.** Ablation study results showcasing the impact of syntax-aware components in SEST.

Model	ACC(%)	self-BLEU	BLEU	CS	WO	PPL
<b>GYAFC data</b>						
SEST	84.1	-	21.1	0.962	0.591	73
SEST w/o Syntax-aware Encoder	83.8	-	20.3	0.957	0.544	83
SEST w/o Syntax-aware Encoder & Classifier	78.7	-	15.6	0.943	0.446	223
<b>Yelp data</b>						
SEST	93.0	57.7	-	0.971	0.778	74
SEST w/o Syntax-aware Encoder	92.6	56.4	-	0.964	0.720	85
SEST w/o Syntax-aware Encoder & Classifier	89.3	49.1	-	0.943	0.697	230

### 5.7. Case Study

Table 8 presents a case study with examples from the GYAFC and Yelp datasets, showcasing the output of SEST and top-performing baselines. The examples illustrate SEST’s ability to effectively transfer style while maintaining content and grammatical correctness.

**Table 8.** Example outputs from SEST and baselines for style transfer tasks. Errors are highlighted.

	From Formal to Informal (GYAFC)	From Positive to Negative (Yelp)
Source	Also, I dislike it when my father is unhappy.	We will definitely come back here!
DualRL	Also <b>i thrilled...</b>	We will not come back here!
DAST	Also, <b>i r it</b> when my father <b>is men!</b>	We will <b>normally joke</b> back here?
PFST	So I miss it when my father is 18.	We will not come back here again.
SEST (ours)	I also hate it when my father is unhappy !!	We will not come back here!

## 6. Conclusion

In this study, we scrutinized the performance of style classifiers employed in prevalent Text Style Transfer (TST) models. Our analysis revealed a significant limitation: these classifiers typically fail to effectively internalize syntactic structures within texts. To address this gap, we introduced the Syntax-Enhanced Style Transfer (SEST) model, an innovative deep learning architecture tailored to integrate syntactic comprehension into the process of style representation learning. Our experimental approach involved rigorous testing across two well-established datasets, where SEST was benchmarked against a range of leading TST models. Through a blend of automated metrics and human evaluations, we established that SEST excels in its domain, outshining existing state-of-the-art approaches. Particularly notable was SEST’s ability to produce sentences in the target style that not only were fluent but also retained the essence of the original content. This capability underscores the effectiveness of SEST in balancing style transformation with content preservation. Looking ahead, we aim to delve deeper into enhancing textual structural representations. Our goal is to refine and integrate these advancements into the SEST framework to elevate its performance in TST tasks further. By continually pushing the boundaries of text style transfer, we aspire to develop more sophisticated and nuanced models capable of handling a broader spectrum of stylistic transformations.

## References

1. Zhiqiang Hu, Roy Ka-Wei Lee, Charu C Aggarwal, and Aston Zhang. Text style transfer: A review and experimental evaluation. *arXiv preprint arXiv:2010.12742*, 2020.
2. Hao Fei, Shengqiong Wu, Yafeng Ren, and Meishan Zhang. Matching structure for dual learning. In *Proceedings of the International Conference on Machine Learning, ICML*, pages 6373–6391, 2022.
3. Tianxiao Shen, Tao Lei, Regina Barzilay, and Tommi Jaakkola. Style transfer from non-parallel text by cross-alignment. In *Advances in neural information processing systems*, pages 6830–6841, 2017.
4. Jake Zhao, Yoon Kim, Kelly Zhang, Alexander M Rush, and Yann LeCun. Adversarially regularized autoencoders. In *35th International Conference on Machine Learning, ICML 2018*, pages 9405–9420. International Machine Learning Society (IMLS), 2018.
5. Hao Fei, Yafeng Ren, Yue Zhang, Donghong Ji, and Xiaohui Liang. Enriching contextualized language model from knowledge graph for biomedical information extraction. *Briefings in Bioinformatics*, 22(3), 2021.
6. Zhenxin Fu, Xiaoye Tan, Nanyun Peng, Dongyan Zhao, and Rui Yan. Style transfer in text: Exploration and evaluation. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.
7. Hao Fei, Yafeng Ren, and Donghong Ji. Boundaries and edges rethinking: An end-to-end neural model for overlapping entity relation extraction. *Information Processing & Management*, 57(6):102311, 2020.
8. Liqun Chen, Shuyang Dai, Chenyang Tao, Haichao Zhang, Zhe Gan, Dinghan Shen, Yizhe Zhang, Guoyin Wang, Ruiyi Zhang, and Lawrence Carin. Adversarial text generation via feature-mover’s distance. In *Advances in Neural Information Processing Systems*, pages 4666–4677, 2018.

9. Zhiting Hu, Zichao Yang, Xiaodan Liang, Ruslan Salakhutdinov, and Eric P Xing. Toward controlled generation of text. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1587–1596. JMLR. org, 2017.
10. Jingye Li, Hao Fei, Jiang Liu, Shengqiong Wu, Meishan Zhang, Chong Teng, Donghong Ji, and Fei Li. Unified named entity recognition as word-word relation classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 10965–10973, 2022.
11. Ning Dai, Jianze Liang, Xipeng Qiu, and Xuan-Jing Huang. Style transformer: Unpaired text style transfer without disentangled latent representation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5997–6007, 2019.
12. Ye Zhang, Nan Ding, and Radu Soricut. Shaped: Shared-private encoder-decoder for text style adaptation. In *Proceedings of NAACL-HLT*, pages 1528–1538, 2018.
13. Jingye Li, Kang Xu, Fei Li, Hao Fei, Yafeng Ren, and Donghong Ji. MRN: A locally and globally mention-based reasoning network for document-level relation extraction. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 1359–1370, 2021.
14. Hao Fei, Shengqiong Wu, Yafeng Ren, Fei Li, and Donghong Ji. Better combine them together! integrating syntactic constituency and dependency representations for semantic role labeling. In *Findings of the Association for Computational Linguistics: ACL/IJCNLP 2021*, pages 549–559, 2021.
15. Shengqiong Wu, Hao Fei, Fei Li, Meishan Zhang, Yijiang Liu, Chong Teng, and Donghong Ji. Mastering the explicit opinion-role interaction: Syntax-aided neural transition system for unified opinion role labeling. In *Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence*, pages 11513–11521, 2022.
16. Wenxuan Shi, Fei Li, Jingye Li, Hao Fei, and Donghong Ji. Effective token graph modeling using a novel labeling strategy for structured sentiment analysis. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4232–4241, 2022.
17. Hao Fei, Yue Zhang, Yafeng Ren, and Donghong Ji. Latent emotion memory for multi-label emotion classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 7692–7699, 2020.
18. Wei Xu, Alan Ritter, Bill Dolan, Ralph Grishman, and Colin Cherry. Paraphrasing for style. In *Proceedings of COLING 2012*, pages 2899–2914, 2012.
19. Hao Fei, Shengqiong Wu, Jingye Li, Bobo Li, Fei Li, Libo Qin, Meishan Zhang, Min Zhang, and Tat-Seng Chua. Lasuie: Unifying information extraction with latent adaptive structure-aware generative language model. In *Proceedings of the Advances in Neural Information Processing Systems, NeurIPS 2022*, pages 15460–15475, 2022.
20. Harsh Jhamtani, Varun Gangal, Eduard Hovy, and Eric Nyberg. Shakespearizing modern language using copy-enriched sequence to sequence models. In *Proceedings of the Workshop on Stylistic Variation*, pages 10–19, 2017.
21. Mingyue Shang, Piji Li, Zhenxin Fu, Lidong Bing, Dongyan Zhao, Shuming Shi, and Rui Yan. Semi-supervised text style transfer: Cross projection in latent space. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 4939–4948, 2019.
22. Shengqiong Wu, Hao Fei, Yafeng Ren, Donghong Ji, and Jingye Li. Learn from syntax: Improving pair-wise aspect and opinion terms extraction with rich syntactic knowledge. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, pages 3957–3963, 2021.
23. Yunli Wang, Yu Wu, Lili Mou, Zhoujun Li, and Wenhan Chao. Harnessing pre-trained neural networks with rules for formality style transfer. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3564–3569, 2019.
24. Hao Fei, Fei Li, Bobo Li, and Donghong Ji. Encoder-decoder based unified semantic role labeling with label-aware syntax. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 12794–12802, 2021.
25. Fengqi Wang, Fei Li, Hao Fei, Jingye Li, Shengqiong Wu, Fangfang Su, Wenxuan Shi, Donghong Ji, and Bo Cai. Entity-centered cross-document relation extraction. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 9871–9881, 2022.
26. Lajanugen Logeswaran, Honglak Lee, and Samy Bengio. Content preserving text generation with attribute controls. In *Advances in Neural Information Processing Systems*, pages 5103–5113, 2018.

27. Shengqiong Wu, Hao Fei, Wei Ji, and Tat-Seng Chua. Cross2StrA: Unpaired cross-lingual image captioning with cross-lingual cross-modal structure-pivoted alignment. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2593–2608, 2023.
28. Di Yin, Shujian Huang, Xin-Yu Dai, and Jiajun Chen. Utilizing non-parallel text for style transfer by making partial comparisons. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 5379–5386. AAAI Press, 2019.
29. Chih-Te Lai, Yi-Te Hong, Hong-You Chen, Chi-Jen Lu, and Shou-De Lin. Multiple text style transfer by using word-level conditional generative adversarial network with two-phase training. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3570–3575, 2019.
30. John Vineet, Lili Mou, Hareesh Bahuleyan, and Olga Vechtomova. Disentangled representation learning for non-parallel text style transfer. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 424–434, 2019.
31. Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.
32. Dianqi Li, Yizhe Zhang, Zhe Gan, Yu Cheng, Chris Brockett, Bill Dolan, and Ming-Ting Sun. Domain adaptive text style transfer. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3295–3304, 2019.
33. Shengqiong Wu, Hao Fei, Leigang Qu, Wei Ji, and Tat-Seng Chua. Next-gpt: Any-to-any multimodal llm, 2023.
34. Yoon Kim. Convolutional neural networks for sentence classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1746–1751, 2014.
35. Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder–decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734, 2014.
36. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008, 2017.
37. Fuli Luo, Peng Li, Jie Zhou, Pengcheng Yang, Baobao Chang, Zhifang Sui, and Xu Sun. A dual reinforcement learning framework for unsupervised text style transfer. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence, IJCAI 2019*, 2019.
38. Zhirui Zhang, Shuo Ren, Shujie Liu, Jianyong Wang, Peng Chen, Mu Li, Ming Zhou, and Enhong Chen. Style transfer as unsupervised machine translation. *arXiv*, pages arXiv–1808, 2018.
39. Sudha Rao and Joel Tetreault. Dear sir or madam, may i introduce the gyafc dataset: Corpus, benchmarks and metrics for formality style transfer. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 129–140, 2018.
40. Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
41. Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24–26, 2017, Conference Track Proceedings*. OpenReview.net, 2017. URL <https://openreview.net/forum?id=SJU4ayYgl>.
42. Diego Marcheggiani and Ivan Titov. Encoding sentences with graph convolutional networks for semantic role labeling. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1506–1515, 2017.
43. Joost Bastings, Ivan Titov, Wilker Aziz, Diego Marcheggiani, and Khalil Sima’an. Graph convolutional encoders for syntax-aware neural machine translation. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1957–1967, 2017.
44. Yuhao Zhang, Yuhui Zhang, Peng Qi, Christopher D. Manning, and Curtis P. Langlotz. Biomedical and clinical english model packages in the stanza python nlp library. *arXiv preprint arXiv:2007.14640*, 2020.
45. Junxian He, Xinyi Wang, Graham Neubig, and Taylor Berg-Kirkpatrick. A probabilistic formulation of unsupervised text style transfer. In *International Conference on Learning Representations (ICLR)*, 2020.

46. Juncen Li, Robin Jia, He He, and Percy Liang. Delete, retrieve, generate: a simple approach to sentiment and style transfer. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1865–1874, 2018.
47. Yixin Liu, Graham Neubig, and John Wieting. On learning text style transfer with direct rewards. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4262–4273, Online, June 2021. Association for Computational Linguistics. URL <https://www.aclweb.org/anthology/2021.naacl-main.337>.
48. Eric Jang, Shixiang Gu, and Ben Poole. Categorical reparametrization with gumbel-softmax. In *Proceedings International Conference on Learning Representations 2017*. OpenReviews.net, April 2017. URL <https://openreview.net/pdf?id=rkE3y85ee>.
49. Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. *OpenAI Blog*, 1(8):9, 2019.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.