experiments, we set $\pi_i = \pi_{labeled,i}$, $\nu = 0.01$, with a dropout rate of 0.1, and other hyperparameters remained unchanged. As shown in Table 4, we compared our framework with the existing state-of-the-art methods ATLOP (**?**), DocuNET (**?**), KD-DocRE (**?**), and SSR-PU (**?**). Our framework achieves the best results as well.

**Effect of Hyperparameters.** Figure 2 shows the effect of hyperparameters on the model under the DocRED and DocRED_ext incomplete labeling settings. (a) shows the effect of the scaling factor $\lambda$, with similar trends under both settings, and $\lambda = 10$ being the best choice. (b) shows the effect of the dropout rate, indicating that the greater the degree of incompleteness in labeling, the greater the dropout rate needed to enhance diversity, but too large a dropout rate will also introduce more noise. (c) shows the effect of $\alpha$, indicating that the model is not sensitive to the choice of $\alpha$, and $\alpha = 1.0$ can be seen as a uniform mixup interpolation between distributions. (d) shows the effect of $\nu$, with similar trends under both settings, and more severe incomplete labeling requires slightly larger mixup strength.

**Effect of Prior Estimation.** Table 5 shows the effect of different prior estimates on the model. It can be seen that our framework is not sensitive to errors in prior estimates, especially in cases where the prior estimate is too small. Even when $\pi_i = \pi_{labeled,i}$, the model still performs well, demonstrating the robustness of our method under errors in prior estimates, which is very helpful for real-world applications.

## Related Work

**Document-Level Relation Extraction.** Previously, effective methods for document-level relation extraction (RE) have mainly been graph-based models and transformer-based models. Graph-based models (**?????**) use graph neural networks to gather entity information for relational inference, while transformer-based methods (**????**) capture long-range dependencies implicitly. Recently, it has been found that there are a large number of false negatives in document-level RE datasets, i.e. incomplete labels (**??**). (**?**) proposed using positive-unlabeled learning to address this problem.

**Positive-Unlabeled Learning.** Positive-unlabeled (PU) learning (**?????**), as a emerging weakly supervised learning paradigm, aims to learn classifiers from positive and unlabeled data, and has gained continuous attention from researchers. PU learning has been widely applied in various tasks, such as text classification (**?**), sentence embedding (**?**), named entity recognition (**??**), knowledge graph completion (**?**), and sentence-level RE (**?**) in the NLP field. (**?**) used PU learning to address the issue of negative samples potentially carrying the same label in contrastive learning.

**Deep Metric Learning.** Our work is inspired by metric learning and mainly falls into two categories: pair-based losses and proxy-based losses. Pair-based methods (**????**) focus on the relationships between individual samples, and contrastive learning can be considered a subset of this approach. Proxy-based methods like Proxy-NCA (**?**) and NormFace (**?**) consider the relationships between proxies and samples, and (**?**) unified the relationship between SoftMax loss and triplet loss, and proposed a new SoftTriplet loss. Proxy-based methods are a type of approach that focuses on improving generalization while keeping training complexity low, although they may not fully utilize the relationships between individual samples.

**Data Augmentation.** Data augmentation is a key factor in deep learning performance and is widely used in many fields (**??**). (**??**) proposed to augment words by randomly inserting and replacing them, while (**?**) augmented the word embeddings directly. (**?**) used simple dropout to augment sentence embeddings for unsupervised contrastive learning. Mixup (**??**) can be considered as another common data augmentation method, where interpolation is used to improve the generalization performance of the model between two samples. It is increasingly used and researched in the NLP (**???**) and the PU learning (**????**) fields. (**?**) proposed a document augmentation dense retrieval framework that uses both methods.

## Conclusion and Future Work

To address document-level RE with incomplete labeling, we propose a positive-unlabeled metric learning framework $P^3M$. First, we combine positive-unlabeled learning with metric learning to learn better representations. Then, we use dropout augmentation to expand the distribution of labeled positive samples. Finally, we use none-class relation embedding as pseudo-negative samples and propose a positive-none-class mixup method to further improve the model's generalization performance. Experiments demonstrate that our method achieve state-of-the-art results in both incomplete and complete labeling scenarios, as well as robustness to prior estimation bias. In the future, we will explore various metric learning losses and data augmentation methods.

## Acknowledgments

This work is funded by National Natural Science Foundation of China (under project No. 62377013).The computation is supported by the ECNU Multifunctional Platform for Innovation (001).

Eligendi nihil quae minima autem soluta quia error delectus aspernatur, corrupti officiis debitis provident eaque sapiente quis deleniti perspiciatis, saepe voluptas molestiae eaque recusandae nemo distinctio expedita hic, magnam commodi a asperiores voluptate.Veniam atque vitae quam ex eum quas, recusandae fuga assumenda eveniet accusantium voluptas quis nesciunt suscipit, veritatis consequatur atque laborum quibusdam fugiat fugit magni incidunt explicabo ratione, fuga quos praesentium nisi laudantium necessitatibus et quae quidem sed impedit corrupti, alias sunt atque deserunt minus molestiae similique exercitationem ut?Possimus quia reprehenderit obcaecati a tempora quidem iure, exercitationem doloremque accusantium beatae.Nisi libero consequatur voluptates iure soluta id earum inventore sed, maxime nostrum a, optio quisquam consectetur fuga laudantium nobis vitae dicta inventore, blanditiis reiciendis repellendus, voluptatibus expedita velit qui aut reiciendis quisquam voluptatem.Error animi mollitia fugit, id alias sint hic asperiores accusantium et corrupti ab quam repellat, cupiditate error ea molestiae ex