

The Insight Circle

1) Abhishek Mishra

2) Anjan Mondal

3) Anuraj Kashyap

4) Akshay Thorave

Topic : Book Genre Classification Using Cover and Title.



Book Genre Classification Using Cover and Title

Accurately classifying books into genres is a challenging task that has applications in areas like personalized book recommendations and library organization. This project explores using a combination of book cover images and titles to train a machine learning model for effective genre classification. By leveraging the power of pre-trained computer vision and natural language processing models, the team was able to achieve state-of-the-art performance.



Leveraging Pre-Trained Models

Cover Feature Extraction

The team utilized the base convolutional layers of the pre-trained Xception model, which was originally trained on the ImageNet dataset, to extract visual features from the book cover images. This allowed the model to leverage the powerful image recognition capabilities learned from a large, diverse dataset.

Title Feature Extraction

For the book titles, the team used the pre-trained GloVe word embedding model to generate feature vectors that capture the semantic and syntactic information of the text. This provided a robust way to incorporate the title information into the classification model.

Model Fusion

By concatenating the cover and title features, the team created a combined feature representation that allowed the final logistic regression model to leverage both visual and textual cues for accurate genre prediction. This fusion of modalities proved to be a key factor in the model's strong performance.



Category:

Dataset and Preprocessing

Data Collection

The dataset used for this project was obtained from the University of Chicago's book dataset, which contains over 2.8 million book covers and associated metadata, including titles and genres. The team focused on a subset of 5 major genres: Children's Books, Cookbooks, Law, Medical Books, and Travel.

Train-Validation-Test Split

The dataset was split into training, validation, and test sets, with the training set comprising approximately 70% of the data, the validation set 20%, and the test set 10%. This allowed for thorough

1

2

Data Preprocessing

The book cover images were resized to the input size required by the Xception model (299x299 pixels) and normalized. The titles were cleaned and tokenized, and the GloVe model was used to generate 300-dimensional feature vectors for each title.

3

Methodology

Cover-Based Classification

The team first evaluated the performance of a model trained solely on the book cover features extracted using the Xception model. This provided a baseline understanding of how much information could be gleaned from the visual appearance of the book covers alone.

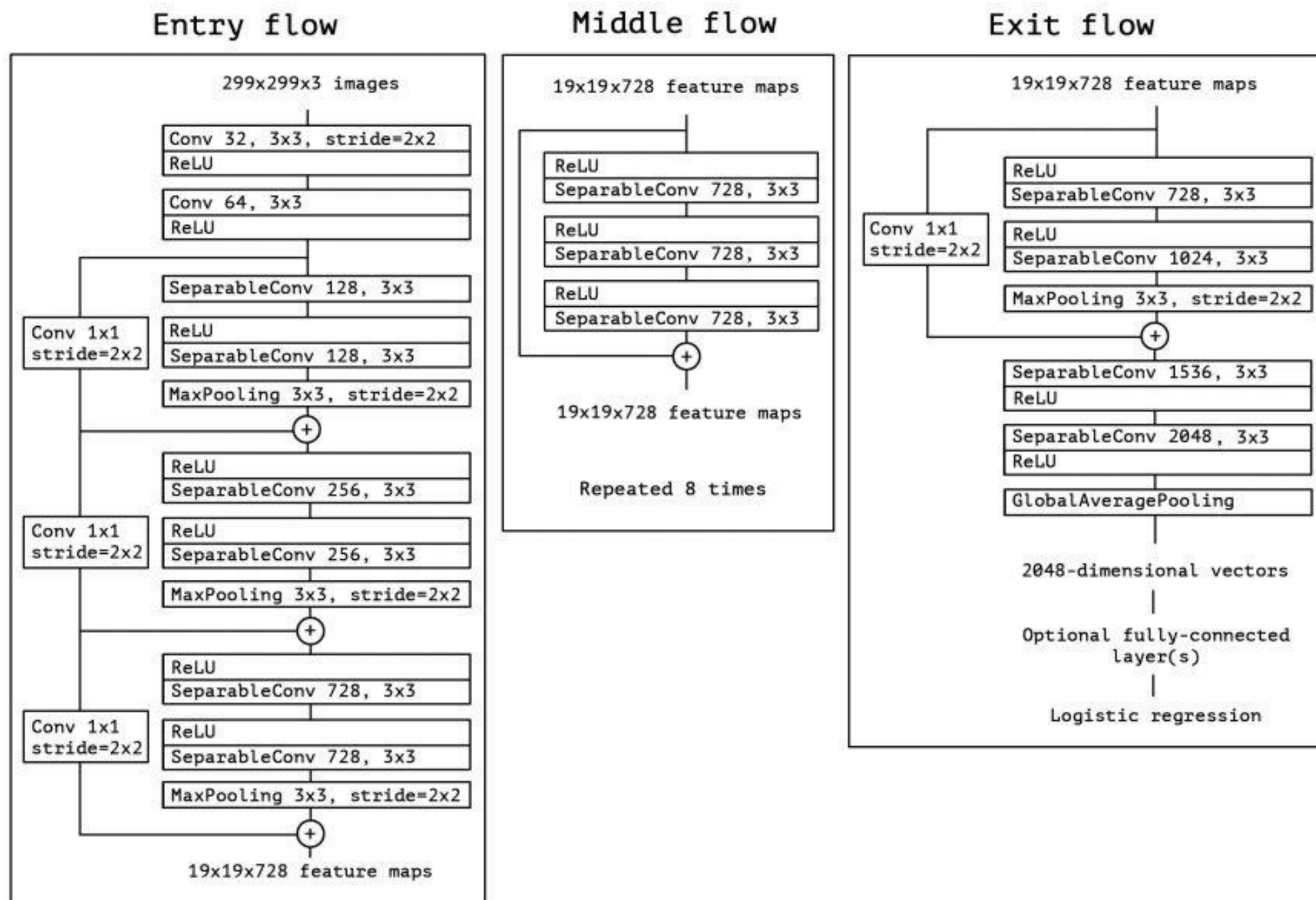
Title-Based Classification

Next, the team trained a model using only the title features generated by the GloVe word embeddings. This allowed them to assess the predictive power of the textual information contained in the book titles.

Combined Model

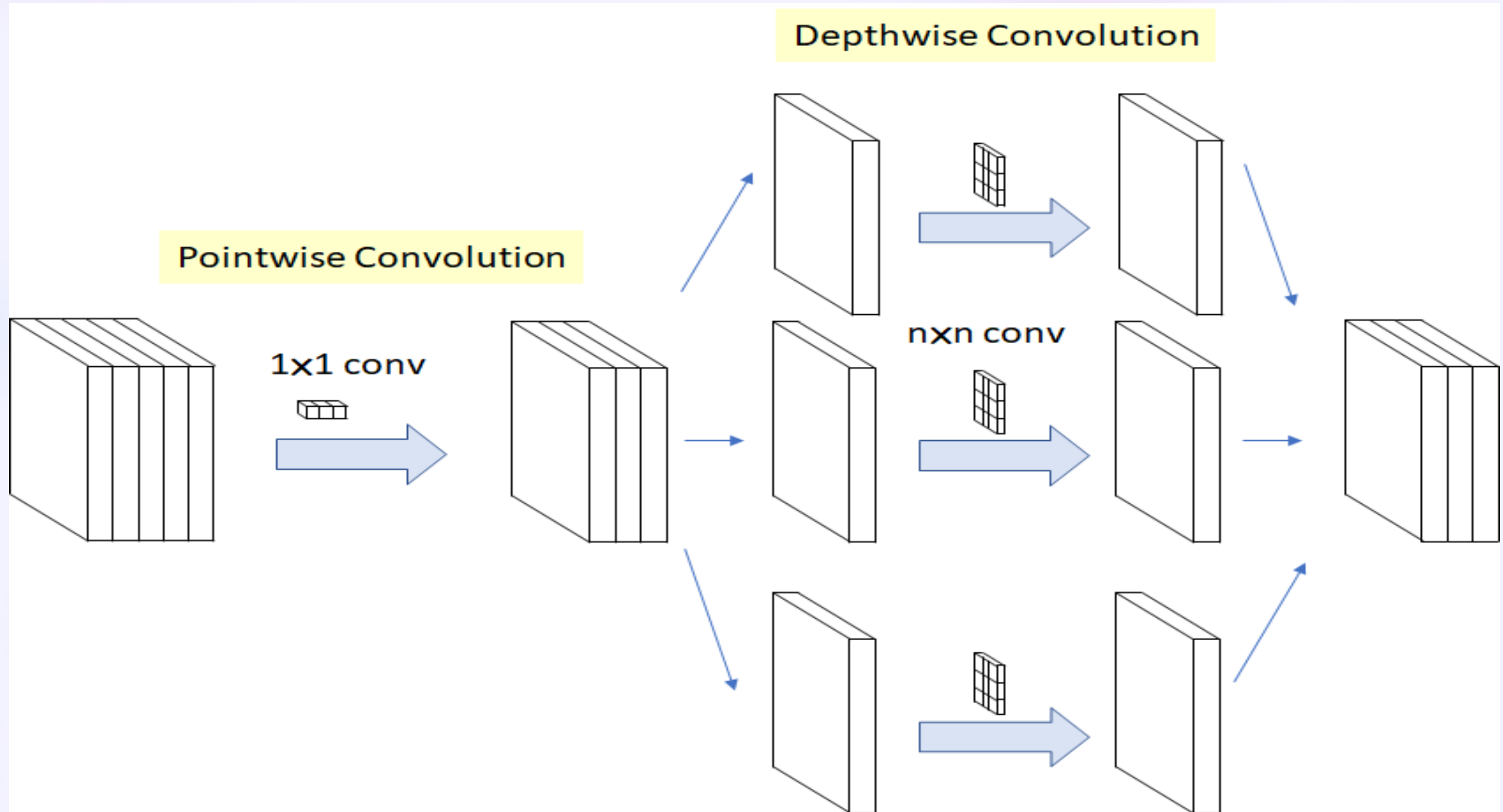
Finally, the team concatenated the cover and title features to create a combined input representation, which was then used to train a logistic regression model for the final genre classification. This fusion of visual and textual cues led to the best overall performance.

Models Used

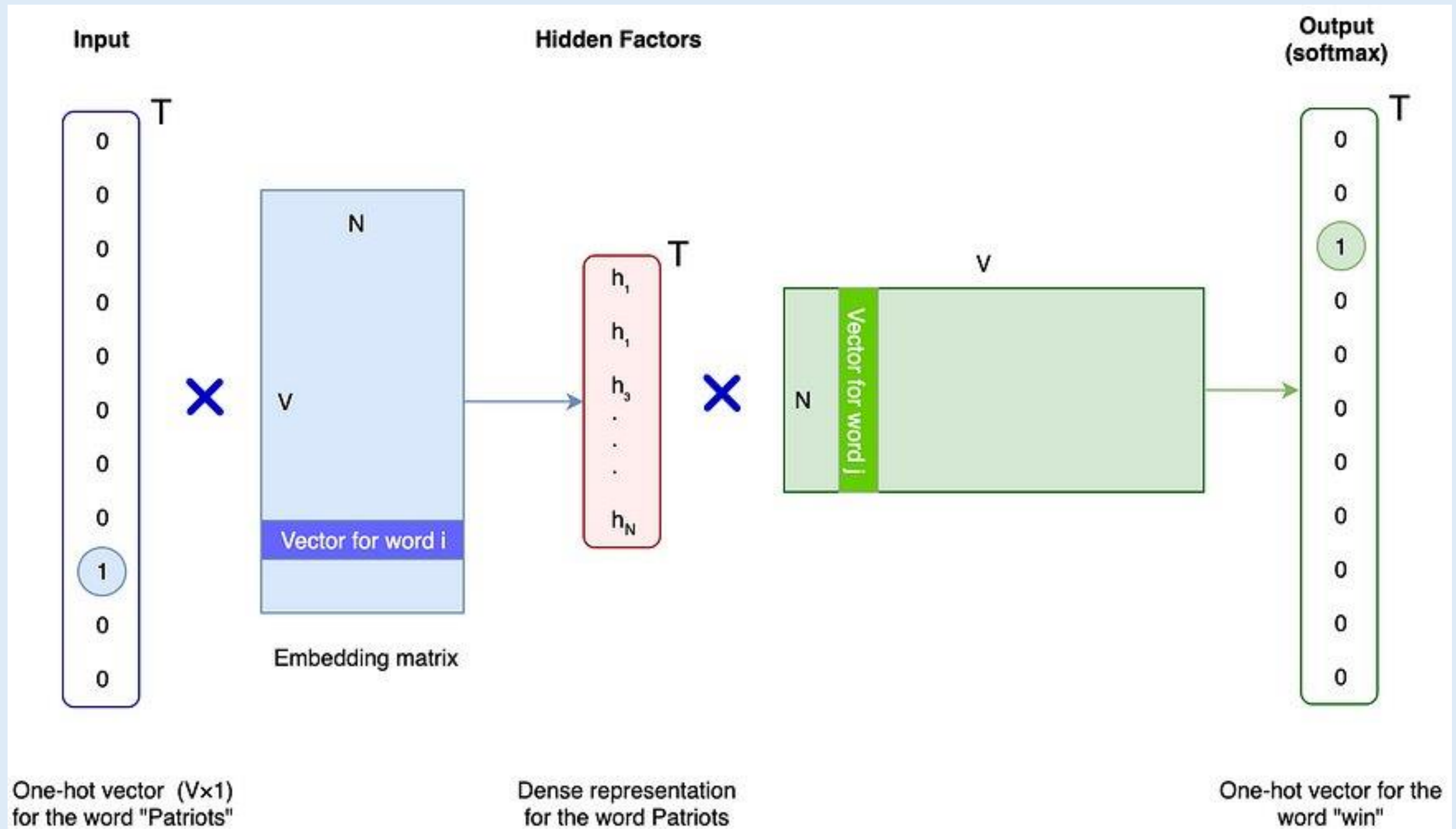


Models Used

Xception



Glove



Results and Insights

0.6704411764705882					
	precision	recall	f1-score	support	
Children	0.75	0.76	0.75	1360	
Cookbooks, Food & Wine	0.82	0.74	0.78	1360	
Law	0.58	0.58	0.58	1360	
Medical Books	0.58	0.61	0.59	1360	
Travel	0.65	0.66	0.66	1360	
avg / total	0.67	0.67	0.67	6800	

XCeption Model Alone

0.8622058823529412					
	precision	recall	f1-score	support	
Children's Books	0.77	0.81	0.79	1360	
Cookbooks, Food & Wine	0.92	0.90	0.91	1360	
Law	0.89	0.86	0.87	1360	
Medical Books	0.89	0.88	0.88	1360	
Travel	0.85	0.87	0.86	1360	
avg / total	0.86	0.86	0.86	6800	

GloVe Model Alone

0.8723529411764706					
	precision	recall	f1-score	support	
Children	0.84	0.85	0.85	1360	
Cookbooks, Food & Wine	0.92	0.91	0.91	1360	
Law	0.88	0.85	0.87	1360	
Medical Books	0.87	0.87	0.87	1360	
Travel	0.86	0.88	0.87	1360	
avg / total	0.87	0.87	0.87	6800	

Combined Model

Results and Insights

1 Accuracy Comparison

- Title-based model outperformed cover-based model with 86% accuracy.
- Combined model utilizing both cover and title features achieved highest accuracy at 87%.

2 Classification Report Analysis

- High recall rates (>90%) observed in "Cookbooks, Food & Wine" and "Children's Books" genres.
- "Law" and "Medical Books" genres exhibited comparatively lower recall scores (85-88%).

3 Insights

- Titles are more informative for genre classification than covers.
- Titles reflect content, while covers focus on visual appeal.

Future Directions

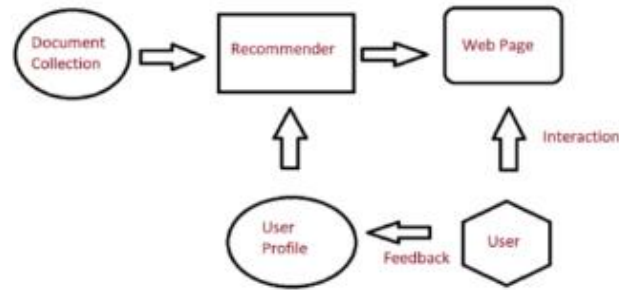


Fig: Recommender System

1

Expanded Genre Taxonomy

Exploring the classification of books across a more diverse set of genres could yield valuable insights and help develop more comprehensive genre understanding models.

2

Multimodal Fusion

Investigating additional modalities beyond cover images and titles, such as book descriptions, reviews, and metadata, could further improve the model's ability to accurately classify book genres.

3

Real-World Applications

Applying these techniques to build intelligent book recommendation systems and streamline library organization could have significant practical impact in the publishing and education sectors.

Dataset and Code Availability



Book Dataset

University of Chicago's book dataset, which is publicly available on GitHub at [\[https://github.com/uchidalab/book-dataset\]](https://github.com/uchidalab/book-dataset).



Project Code

The code for this project is at [\[https://github.com/anjanmondal/Book_Genre\]](https://github.com/anjanmondal/Book_Genre).



Intermediate Files

Due to file size constraints, the pre-trained GloVe vectors and other intermediate feature files used in the project are not included in the GitHub repository. However, these files can be generated using the provided code.

Conclusion

Effective Genre Classification

- Leveraged pre-trained computer vision and natural language processing models.
- Fused visual and textual features for book genre classification.
- Achieved 87% accuracy on the combined model

Practical Applications

- Techniques applicable in real-world scenarios like personalized book recommendations, library organization, and publishing content analysis.
- Dataset and code availability fosters further research and development in the field.

Future Research Directions

- Avenues for future research: expanding genre taxonomy, exploring additional modalities, and deploying the system in practical applications.
- Continued advancements hold potential for enhancing understanding, organization, and interaction with literary content in the digital age.

References

- [1] Chollet, F. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- [2] Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP).
- [3] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. The Journal of Machine Learning Research, 12, 2825-2830.
- [4] Uchidalab. (2017). Book Dataset. Retrieved from <https://github.com/uchidalab/book-dataset>