

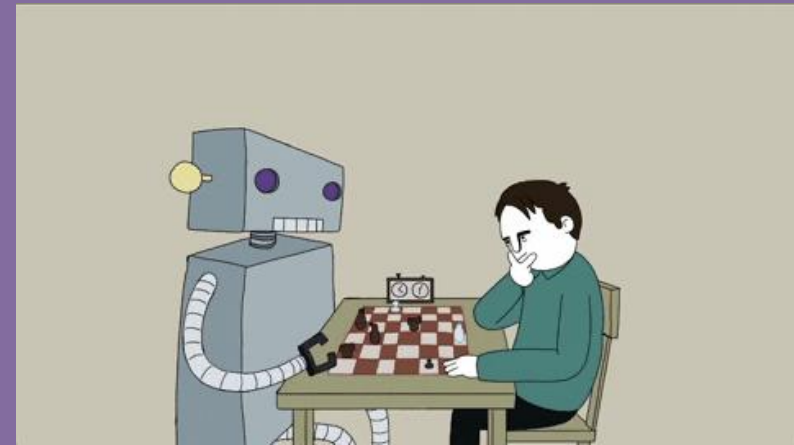
CSE4006

DEEP LEARNING

Dr K G Suma

Associate Professor

School of Computer Science and Engineering



Module No. 6

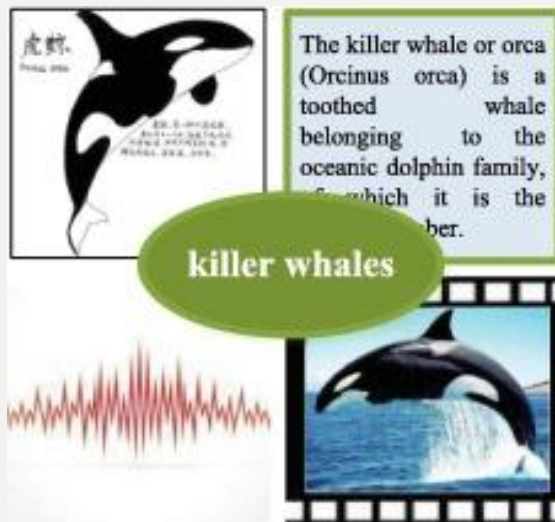
VAEs and GANS

9 Hours

- Variational Autoencoders
 - Generative Adversarial Networks
 - Multi-task Deep Learning
 - Multi-view Deep Learning
-
- Various Applications - speech, text, image and video

Multi-view deep learning

- Multi-view deep learning leverages multiple distinct perspectives or "views" of the same data to improve model performance, generalization, and robustness, particularly in scenarios with limited labeled data or when dealing with complex relationships between data.
- It's essentially a form of deep learning where a model is trained on data represented in several different ways, and then uses the combined information to make predictions.



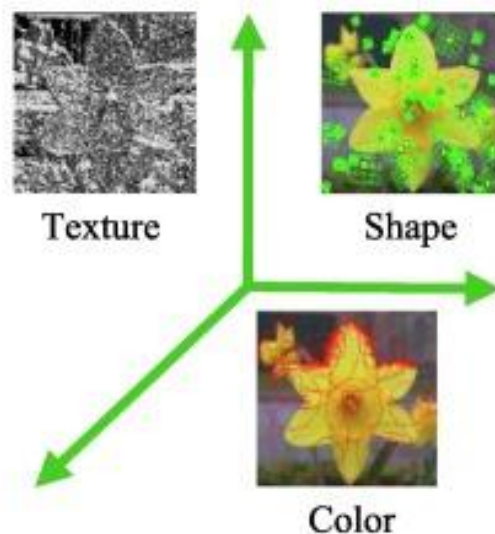
(a) An object is described by text, video, audio



(b) A news is reported by different languages



(c) A product can be represented by multi-view graphs



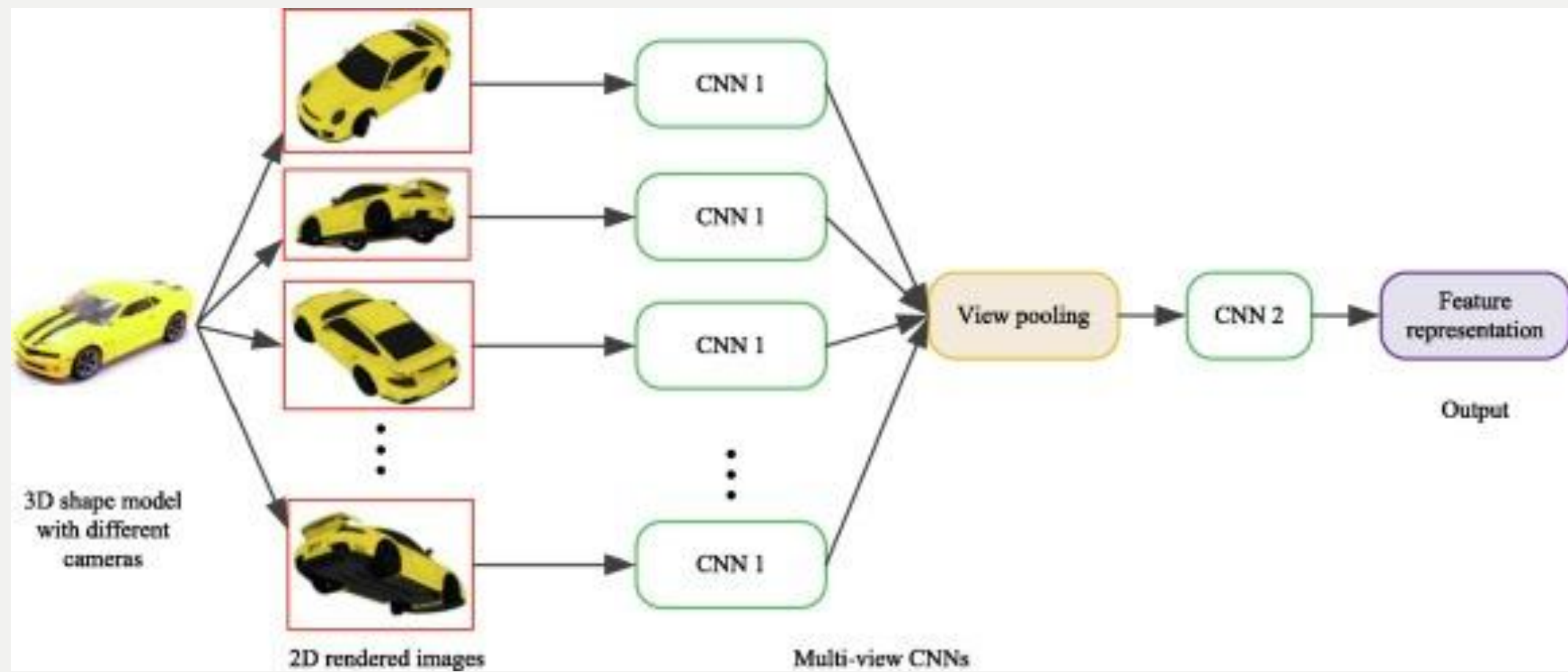
(d) An image is depicted by heterogonous features



(e) A social image along with user tags



(f) One action can be captured from different camera viewpoints



Key Concepts:

Multiple Views: Data can be represented in various ways, such as images, text, audio, or different feature sets. Each representation is considered a "view".

Complementarity: Different views often provide complementary information about the data, meaning they capture different aspects or features that can be beneficial for learning.

Generalization: By considering multiple perspectives, the model can learn more robust and generalized representations, which can better handle unseen data or variations in the input.

Overfitting: Multi-view learning can help mitigate overfitting, which is a common problem in deep learning, particularly when the training data is limited.

Benefits of Multi-view Deep Learning:

Improved Performance: By incorporating information from multiple sources, the model can achieve higher accuracy and better performance on various tasks.

Enhanced Robustness: Multi-view learning can make models less sensitive to noise or variations in the input, leading to more robust and reliable predictions.

Better Generalization: The model can learn more generalizable patterns and relationships between data, leading to better performance on unseen data.

Handling Complex Relationships: It can effectively address complex relationships between different aspects of data, which can be challenging for single-view models.

Applications:

Computer Vision: Combining different visual modalities (e.g., images, depth maps) for object recognition or scene understanding.

Natural Language Processing: Integrating text, audio, and visual information for tasks like speech recognition or multimodal dialogue systems.

Bioinformatics: Analyzing genomic, proteomic, and imaging data for disease prediction or drug discovery.

Recommender Systems: Leveraging user behavior, demographic data, and content information for personalized recommendations.

Challenges:

View Inconsistency: Different views may not always be perfectly aligned or consistent, which can make it difficult to combine them effectively.

View Complementarity: Identifying the most complementary views and how to best combine them can be challenging.

Dimensionality: Combining multiple views can increase the dimensionality of the input data, which can require more complex models and computational resources.

Scalability: Training models on large datasets with many views can be computationally expensive and time-consuming.

Conclusion

Multi-view deep learning is a powerful technique for improving model performance and generalization by leveraging complementary information from multiple data perspectives. It offers a promising approach for tackling complex real-world problems across various domains.