# Chapter 8

# Conclusion and Future Work

**Chapter 8:**

**Conclusion and Future Work**

**8.1 Conclusion**

Lecture Video and PowerPoint presentations are the most multimedia commonly used in e-learning. Extracting Micro LOs from this multimedia are becoming necessary for personalized/adaptive learning. In this thesis, we presented a framework that enables the extraction and annotation of MLOs out of PPT-Presentation lectures, PPT-based Video lectures with availability of the original PPT-presentation, and PPT-based Video lectures without associated PPT Presentation. For PPT-Presentation, a standard PPT lecture template is suggested to simplify the identification and extraction process and for those existing non-standard-format PPT lectures, a rule-based algorithm was developed for automatically converting those non-standard PPT lectures into the suggested standard format.

for the second type of lecture - the PPT-based Video lectures with availability of the original PPT-presentation—the PPT presentation is separately manipulated as before to segment it into its constituent MLOs, and then the video is segmented accordingly by measuring the similarity between the video frames and the PPT MLOs. The result of this second segmentation algorithm is a set of PPT-based MLO video lectures with availability of their corresponding PPT-presentations, similar to the type of the input lecture.

for the last type of lectures- namely, PPT-based Video lectures without associated PPT Presentation—a series of deep learning models were designed to extract Presentation slides out of the video lecture. The first model is concerned with classifying video frames into Slides and Non-Slides frames; the second model is concerned with classifying slide-frames into Cover, Agenda, Separator and Explanation slides; the third and final model is concerned with identifying the contents of the slides into Title and Content. For training and evaluating those deep learning models we have created three different datasets, the first dataset contains 100000 images (50000 slide image and 50000 Non slide images), the second dataset contains 50000 images (Agenda: 4500, Cover: 7800, Separator: 13000, Explanation: 25000), the third dataset contains 10000 images Annotated with (Title: 9495 and Text: 6122). the accuracies of three models, the first model accuracy is (97%), the second model is (88%), the third model is (93%). After extracting the presentation slides, a PPT presentation is created using those slides which is then fed to the previous PPT MLOs Extractor. A corrective rule-based algorithm was developed to improve the accuracy of the resulted PPT presentations to become 99%. Afterwards, the video lecture is segmented into its MLOs corresponding to the automatically generated PPT MLOs. At the end, a rule-based algorithm is designed to annotate the auto-generated MLOs with the appropriate metadata such as (introduction, theory proving, example, or experiment, etc.).

**8.2 Suggestions for Future Work**

The following aspects could be explored in future work:

1. More investigation of Nonstandard presentations discrepancies.
2. Enhancing accuracies of proposed deep learning models.
3. Remove noise from Video MLO segment (end and start)
4. Segment Video MLO at the beginning of sentence.
5. Extract MLOs out of Video Lectures (Non-Based Presentation)