STATS 402 - Interdisciplinary Data Analysis <Building Change Detection using CNN Model on Aerial Images: A Comparative Analysis of 2012 and 2016 Datasets> Milestone Report: Stage 3

1st Jingheng Huan
Duke Kunshan University
Beijing, China
jh730@duke.edu

2nd Luyao Wang
Duke Kunshan University
Suzhou, China
lw337@duke.edu

3rd Fanbin Xu
Duke Kunshan University
Chengdu, China
fx31@duke.edu

***Abstract*—In this milestone report, we will discuss the current status of our satellite image analysis project, including detailed techniques adopted, any adjustments made to our technical route, our existing results, and our plan for finalizing the final report. We will also highlight the applications of our project in real-world scenarios.**

I. Current Status and Techniques Adopted

A. Technical Adjustments and Rationale

Our team has made some adjustments to the previously implemented change detection model. The initial model utilized a ResNet50 FCN (fully convolutional network) with shared weights for both 2012 and 2016 images as a feature extractor. A simple FCN was added after the output of the Siamese ResNet feature extractor to predict the difference image. The model took two satellite images from different time periods as input and predicted the difference image, i.e., the ground truth change label.

However, during the model tuning process, we observed that the accuracy was not improving, and the binary cross entropy remained around 0.69. This indicated that the model failed to learn any significant information from the input data.

We identified the root cause of this issue after discussing it with our instructor. The original model did not utilize the 2012 and 2016 building labels, only the ground truth change label. Consequently, the model was unable to recognize building features, only the changed parts of the buildings. To enable the deep learning model to effectively predict building changes, we needed to incorporate the building labels from both 2012 and 2016 into the network.

B. Updated Model and Techniques

Considering these findings, we made the following adjustments to our approach:

We trained two separate semantic segmentation models for the 2012 and 2016 images. These models take satellite images as input and predict the building labels.

We replaced the input of the Siamese network from two satellite images to the building labels predicted by the semantic segmentation models.

C. Existing Results

At this stage of our project, we are yet to obtain definitive quantitative results from our updated model. However, preliminary tests and observations have shown promising indications of the model's enhanced learning abilities, particularly in its capacity to identify relevant features more efficiently, which we expect to contribute towards an improvement in overall accuracy.

Fig. 1 and 2 illustrate some early outcomes from our revised model's application. For the purpose of simplification and enhanced clarity, we have presented the labels in these figures, while in practice, the model operates using satellite images as input for the semantic segmentation process.

The images suggest that our revised Siamese network, which now incorporates input from two distinct semantic segmentation models, has successfully adapted to discern the variances between two satellite images when the building structures and their changes are significantly pronounced.

As depicted in Fig. 1, in instances where the changes are notably substantial or the building blocks are relatively large, the model is proficient in identifying and learning the change patterns. This is indicative of the model's potential for applications involving large-scale urban changes, infrastructural developments, and significant environmental transformations.

However, as shown in Fig. 2, the model currently exhibits limitations in identifying changes when the building blocks within the satellite images are smaller or the alterations are less pronounced. This suggests that the model might struggle to capture more subtle or minor changes in building structures, which can often be crucial in fields such as historical preservation, precise urban planning, or detailed environmental impact assessments.

Addressing this limitation will be a key focus in our ongoing model refinement process. By improving the model's ability to recognize and learn from smaller patterns of change, we aim to enhance its precision and broaden its applicability. We anticipate that the continued refinement of our model will yield improved quantitative results, which we will present in detail in our final report. While we acknowledge that our model is still in the developmental stage, these preliminary results provide valuable insights into the model's potential strengths and areas for improvement, guiding our future efforts towards achieving a more robust and accurate change detection tool.

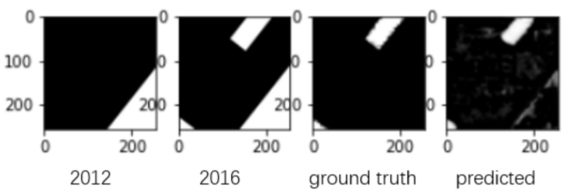

Fig. 1. Successful prediction by the Siamese network with input from two semantic segmentation models.

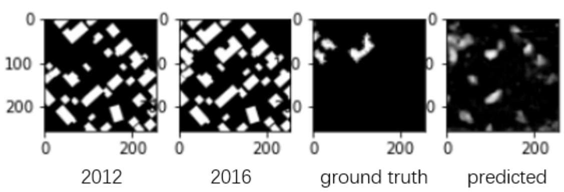

Fig. 2. Unsuccessful prediction by the Siamese network with input from two semantic segmentation models.

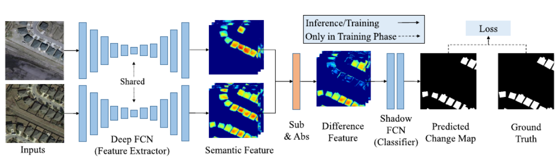

Fig. 3. Illustration of Deep Learning Network with Siamese network for Change Detection [1].

II. Plan for the Final Outcome

A. Model Refinement and Evaluation

Hyperparameter Tuning: One of our primary tasks in the coming weeks will be to conduct a comprehensive hyperparameter tuning process. This involves systematically adjusting the model parameters, such as learning rate, batch size, the number of hidden layers, and dropout rate, to identify the optimal configuration that results in the highest predictive performance. We plan to use strategies such as grid search and random search, which are both common practices in hyperparameter optimization.

Refinement of Methodology: The current Siamese network, incorporating input from two semantic segmentation models, demonstrates limited proficiency in detecting transformations pertaining to smaller building blocks within satellite images. We hypothesize this deficiency originates from alignment discrepancies. For larger buildings, minor offsets in alignment are less impactful due to their relative insignificance to the overall structure. However, for smaller structures, these offsets can result in spatial deviations of 20%-30%.

Given the impracticality of manually labeling the dataset to ameliorate this issue, alternative solutions must be explored. One proposition involves cropping larger segments from the original images and resizing them to 256×256 pixels. However, this method fails to fundamentally resolve the alignment problem. A second suggestion entails modifying the loss function. Our current model employs pixelwise binary cross entropy for change detection, which may not be optimal as it uniformly weighs all pixels, including those at the edges of building blocks that are most susceptible to alignment-induced errors.

With sufficient time, we plan to investigate the potential of attention-based mechanisms within the Siamese network. Theoretically, such mechanisms should enhance the model's performance with smaller building blocks by enabling the learning of spatial-temporal dependencies between pixels. Prior research [2] has indicated promise in this approach, as shown in Fig. 4, prompting us to consider the adoption of an attention-based Siamese Fully Convolutional Network (FCN) model.

Model Evaluation: Concurrent with hyperparameter tuning, we will evaluate our model's performance rigorously. We will use multiple metrics to ensure we capture a holistic view of our model's performance:

a. Accuracy: This is a fundamental measure that shows the proportion of correct predictions made by the model out of all predictions. However, it may not be informative in the case of class imbalance.

b. Precision: This measure will show us how many of the buildings that our model predicted as changed were actually changed. High precision implies a low false positive rate.

c. Recall (or Sensitivity): This will tell us what proportion of actual changes were identified correctly by our model. High recall implies a low false negative rate.

d. F1-score: Given that precision and recall are equally important for our task, the F1-score will help us balance these two measures. It is the harmonic mean of precision and recall and gives a better measure of the incorrectly classified cases than the accuracy metric.

Model Validation: To prevent overfitting and ensure the generalizability of our model, we will also use cross-validation techniques. This involves partitioning our data set into subsets, training the model on some subsets (training set), and validating the model on the remaining subsets (validation set). This process is repeated several times, with each subset serving as the validation set at least once. Cross-validation will help us ensure that our model performs well on unseen data.

B. Real-world Applications

Urban Planning and Development: Our project can play a pivotal role in urban planning and development initiatives.

By providing a detailed and accurate analysis of changes in building structures over time, our model can help urban planners and government authorities make more informed decisions. For example, the model can identify areas of rapid urbanization, where infrastructure development is outpacing planning. This information can be used to prioritize resources, make zoning decisions, and plan future urban development projects. Additionally, it can be used to monitor compliance with building regulations and urban development plans.

Disaster Management: The model can also be invaluable in disaster management. By identifying changes in structures, the model can help authorities assess the impact of natural disasters such as earthquakes, hurricanes, and floods more quickly and accurately. This can speed up response times, help in the efficient allocation of resources, and ultimately save lives. Post-disaster, the model can be used to monitor the reconstruction progress and ensure the efficient use of resources.

Infrastructure Development: In the field of infrastructure development, our model can be used to monitor the progress of large-scale construction projects. By comparing images over time, it can provide an objective assessment of whether the project is progressing as planned. This can help in early identification of potential delays or issues, allowing for timely intervention.

Environmental Conservation: Our model can also play a role in environmental conservation efforts. By monitoring changes in urban areas, the model can help identify encroachments into protected natural areas or changes in land use that could have environmental impacts. This can provide valuable data for conservation organizations and government bodies tasked with protecting the environment.

Historical Preservation: Finally, our model can be used for historical preservation. By monitoring changes in historically significant areas, our model can help identify illegal construction or alterations to protected structures. This can assist authorities in preserving cultural heritage and maintaining historical integrity.

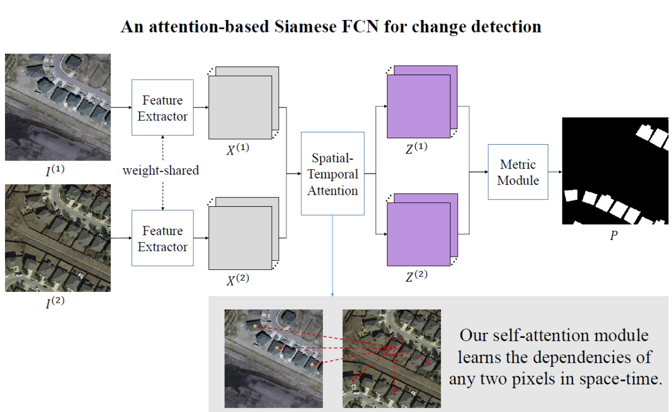

Fig. 4. Illustration of attention-based Siamese FCN network for Change Detection [2].

III. Conclusion

In conclusion, our project to develop a deep learning model for detecting changes in building structures from satellite images has seen significant progress. We have critically evaluated our initial approach and made essential adjustments to improve our model's learning effectiveness. Our revised model, which now incorporates building labels from 2012 and 2016, shows promising initial results, demonstrating an improved learning capacity.

Our upcoming plans involve rigorous hyperparameter tuning, methodology refinement, multi-metric model evaluation, and cross-validation, all aimed at refining our model and ensuring it is robust, accurate, and generalizable.

The potential real-world applications of our project are substantial. They span across urban planning and development, disaster management, infrastructure development, environmental conservation, and historical preservation. By providing accurate, timely, and objective assessments of changes in building structures, our model can inform and improve decision-making in these fields.

As we move towards the final stages of our project, we remain committed to delivering a tool that can make a significant impact in these areas. We anticipate that our final report will present a refined model that effectively predicts changes in building structures and provides valuable insights for various real-world applications.

Our team acknowledges the challenges that lie ahead in finalizing our model and report, but we are confident in our ability to overcome them and deliver a product that meets our objectives and contributes positively to society.

References

[1] H. Chen, W. Li and Z. Shi, "Adversarial Instance Augmentation for Building Change Detection in Remote Sensing Images," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-16, 2022, Art no. 5603216, doi: 10.1109/TGRS.2021.3066802.

[2] H. Chen and Z. Shi, "A Spatial-Temporal Attention-Based Method and a New Dataset for Remote Sensing Image Change Detection," Remote Sensing, vol. 12, no. 10, p. 1662, May 2020, doi: https://doi.org/10.3390/rs12101662.