2021

Ship Deck Segmentation in Engineering Document Using Generative Adversarial Networks

Mohammad Shahab Uddin
Old Dominion University

Raphael Pamie-George
Old Dominion University

Daron Wilkins
Old Dominion University

Andres Sousa Poza
Old Dominion University

Mustafa Canan

See next page for additional authors

Original Publication Citation

Uddin, M. S. U., Pamie-George, R., Wilkins, D., Sousa-Poza, A., Canan, M., Kovacic, S., & Li, J. (2022). In 2022 IEEE World AI IoT Congress (AIIoT) (pp. 207-212). IEEE. http://dx.doi.org/10.1109/AIIoT54504.2022.9817355

Authors
Mohammad Shahab Uddin, Raphael Pamie-George, Daron Wilkins, Andres Sousa Poza, Mustafa Canan, Samuel Kovacic, and Jiang Li

Ship Deck Segmentation In Engineering Document Using Generative Adversarial Networks

Mohammad Shahab Uddin [¥], Raphael Pamie-George[¥], Daron Wilkins[¥], AndresSousa-Poza[°], Mustafa Canan[♦], Samuel Kovacic[°], Jiang Li[¥]
[¥]*Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, Virginia,USA.*
[°]*Departmentt of Engineering Management & Systems Engineering, Old Dominion University, Norfolk, Virginia,USA.*
[♦]*Department of Information Sciences, Naval Postgraduate School, Monterey, California,USA.*
Emails: muddi003@odu.edu; rpami001@odu.edu; dwilk003@odu.edu; asousapo@odu.edu; anthony.canan@nps.navy.mil; skovacic@odu.edu; jli@odu.edu

***Abstract*— Generative adversarial networks (GANs) have become very popular in recent years. GANs have proved to be successful in different computer vision tasks including image-translation, image super-resolution etc. In this paper, we have used GAN models for ship deck segmentation. We have used 2D scanned raster images of ship decks provided by US Navy Military Sealift Command (MSC) to extract necessary information including ship walls, objects etc. Our segmentation results will be helpful to get vector and 3D image of a ship that can be later used for maintenance of the ship. We applied the trained models to engineering documents provided by MSC and obtained very promising results, demonstrating that GANs can be potentially good candidates for this research area.**

Keywords—GAN, Pix2Pix GAN, ship deck segmentation,

I. Introduction

3D view of a ship is very important for proper maintenance of the ship. If 3D information of a ship is available, then we can easily monitor the condition of the ship and plan the maintenance accordingly. It will increase the life cycle of a ship. The current main challenge is that for most of the ship built decades ago, MSC only has 2D scanned images or engineering documents of ship decks. Moreover, there is no clean deck images related to the ship that can be used to train machine learning models for deck segmentation. In this study, we created our own dataset to tackle these challenges and trained GAN models based with the dataset for ship deck segmentation, which will be helpful for generating vector and 3D images for the ships in future.

Many researchers have performed semantic segmentation for building floor plans and many datasets are available for performing research related to building floor segmentation. [1-6]. However, there is not much work on semantic segmentation for ships, making our research for ship deck segmentation challenging. Moreover, ships typically have more complex structures than buildings. Therefore, the architecture plans of the ships are far more complex than the plans of buildings. When we want to clean a ship deck plan, segmentation is a good candidate. Segmentation helps detect and separate various objects in raster images. During the cleansing step, we identify unnecessary objects from deck plan and remove them. To achieve the objective, we applied segmentation methods to ship deck plan images to segment out main plans and eliminate unnecessary information in the image. We trained GAN models with limited dataset for segmentation.

Our work potentially can help proper maintenance of ships and prolong their life cycle. With the clean deck plans, it is possible to convert these traditional engineering documents to vector images and generate 3D models for ships to build digital twin. Digital twin will help improve maintenance properly. During the earlier days, we used to have more crew members than today for ship maintenance. At present, it is more challenging to perform maintenance of ship with the limited number of crew members. Moreover, this will also provide safer working environment for the crew members. Our research will greatly contribute to this area to ensure safety and proper operation of ships operating today. Our contributions are:

1. We have evaluated the use of the generative adversarial networks [7] for ship deck segmentation in traditional engineering documents. We customized the Pix2pix GAN [8] for our experiments.

2. We have created a dataset for ship deck segmentation task. We have included 50 paired images in the dataset from the AutoCAD platform [9] of ship deck plans.

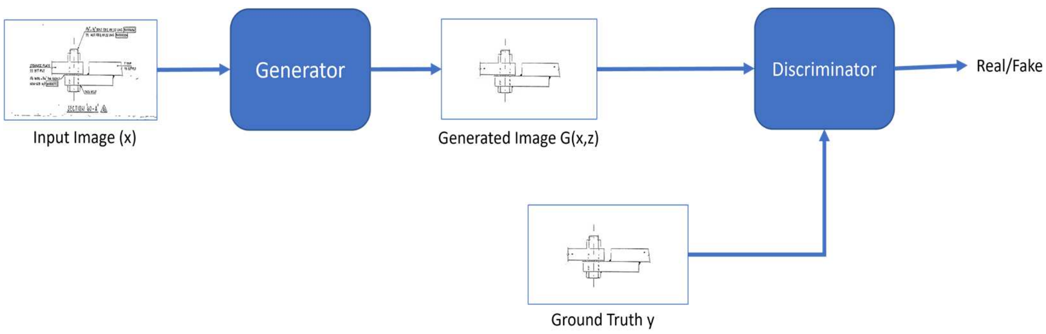


Fig. 1. Overall architecture of the Pix2Pix GAN.

2022 IEEE World AI IoT Congress (AIIoT) | 978-1-6654-8453-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/AIIOT54504.2022.9817355

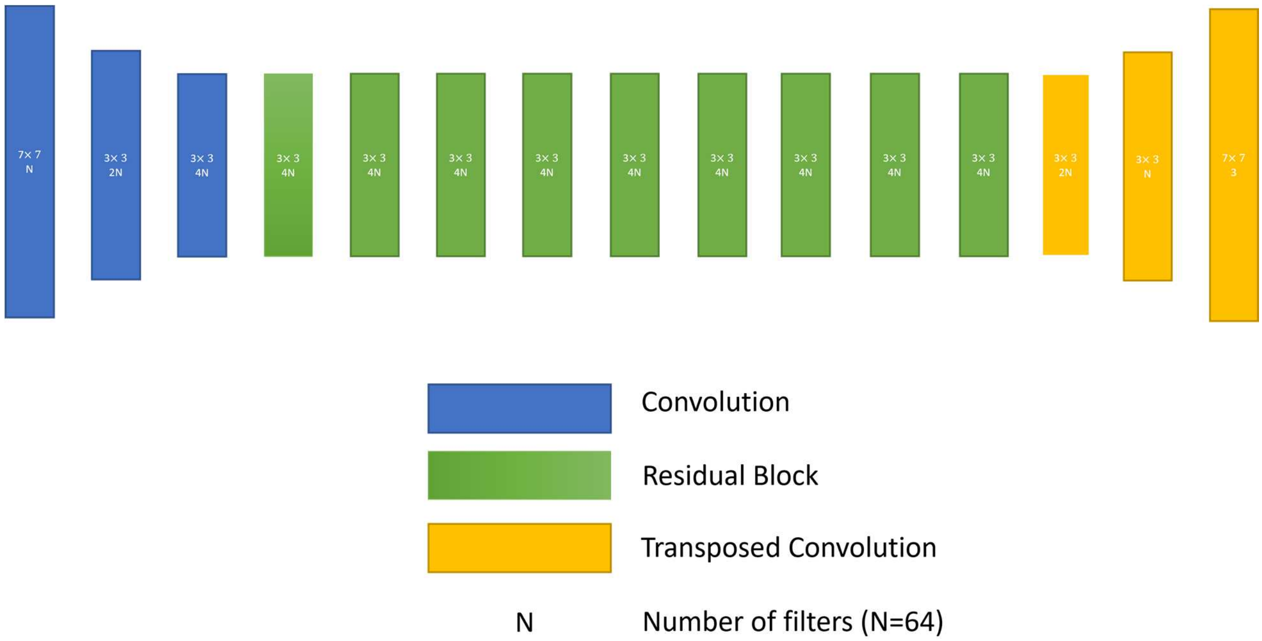


Fig.2 Internal Architecture of generator G in Pix2Pix GAN model.

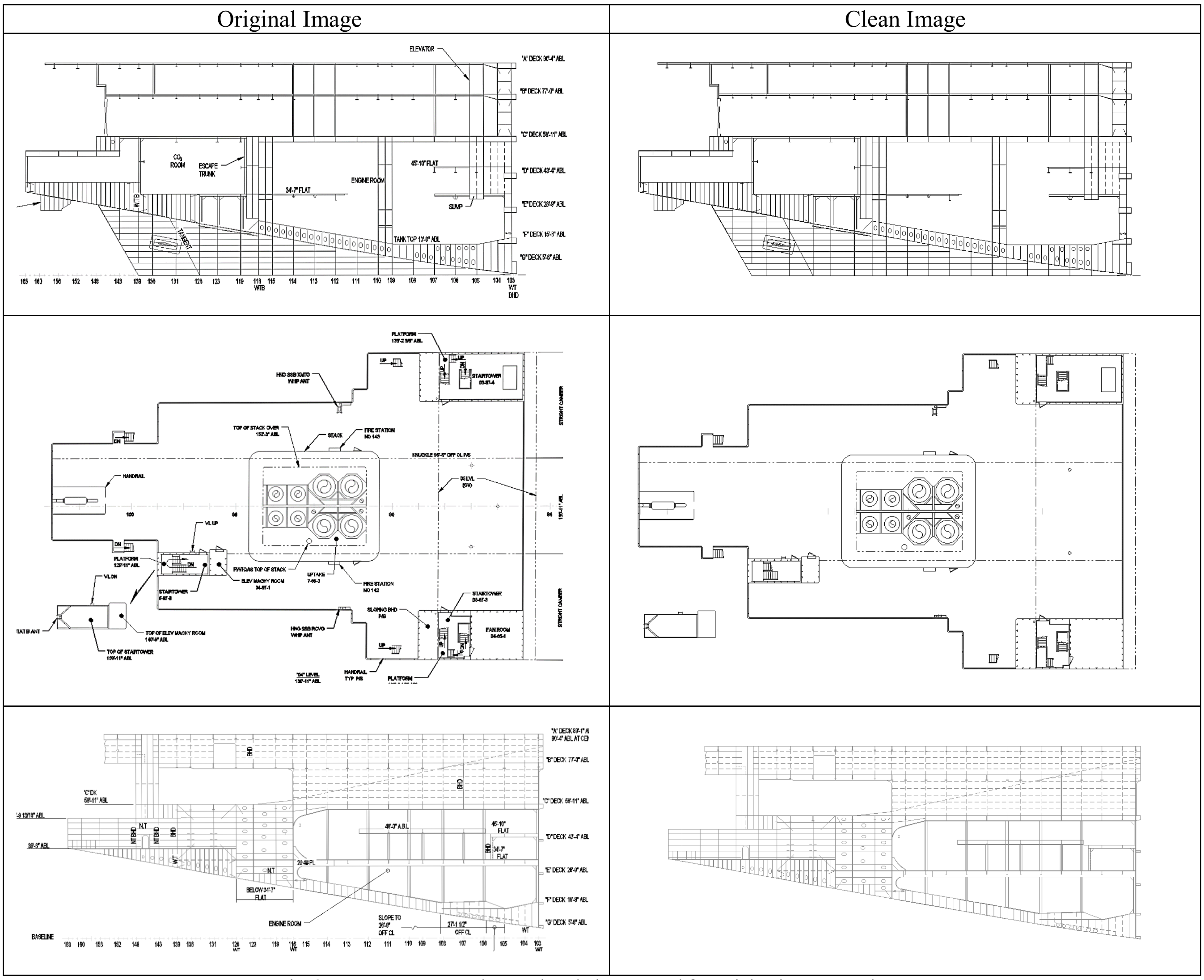


Fig. 3 AutoCAD generated ground truth dataset used for training in our experiments.

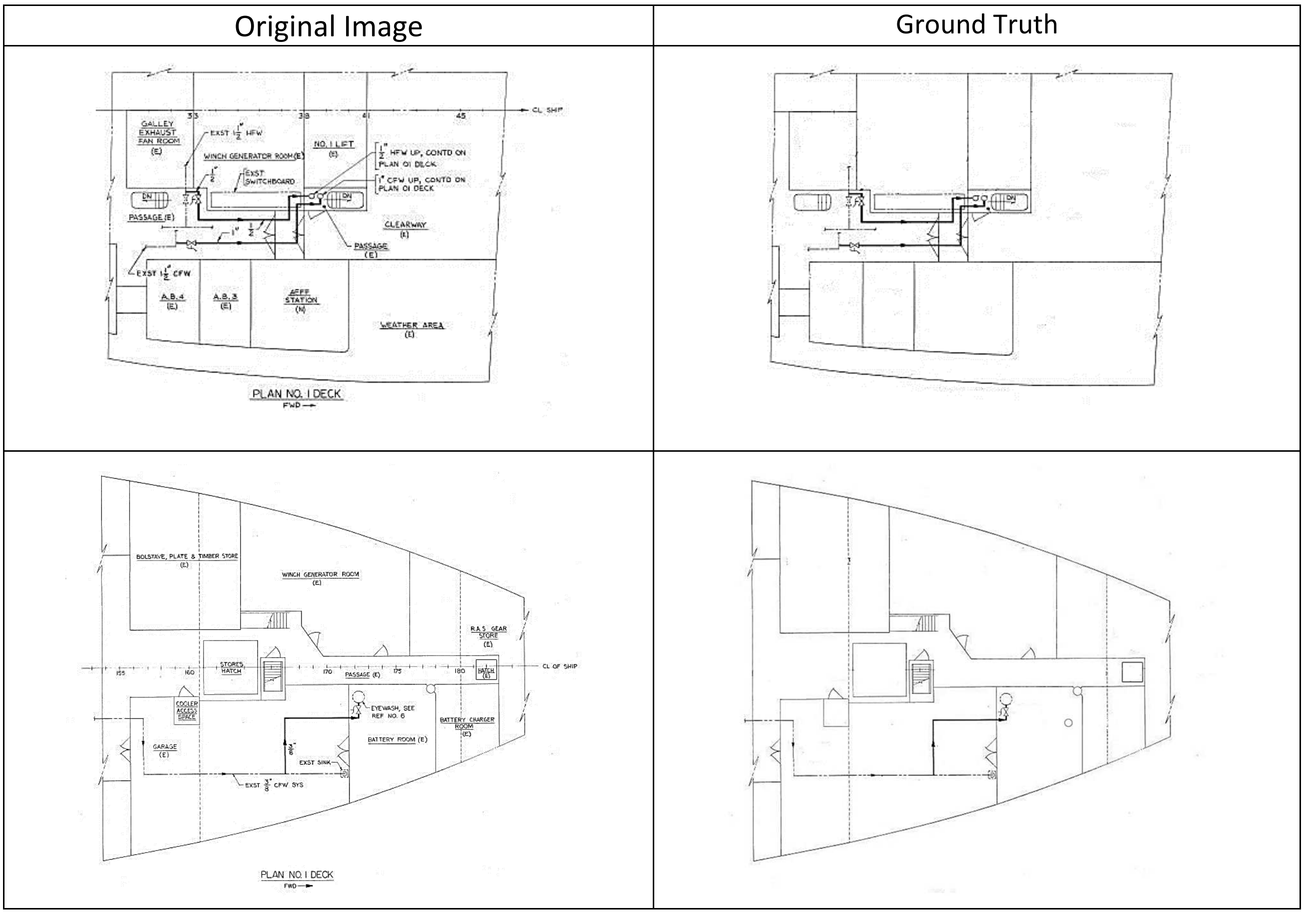


Fig. 4. MSC engineering documents as scanned images of ship deck plans.

II. RELATED WORKS

There is not much work related to ship deck segmentation to the best of our knowledge. However, many work on building floor plan segmentation had been conducted in literature using different methods. For instance, a Fuzzy rule-based system was proposed in [1] for semantic segmentation of building floor plan. In [2], researchers did line segment vectorization to generate 3D models from 2D CAD drawings. Convolutional neural network (CNN) based methods were also used for building floor plan segmentation as illustrated in [3]. The authors used CNN to detect objects from floor plan, and incorporated attention mechanism to perform room boundary prediction and room type prediction simultaneously. In [4], the authors used deep neural networks to extract room structures information from floor plan and generated 3D reconstruction. They also created a dataset for building floor plan segmentation consisting of 7000 floor plans. Researchers in [5] used floor plans and related photos to generate 3D mesh models for buildings. They utilized graph convolutional neural network for segmentation. Moreover, optical character recognition (OCR) was used in this research area as well. In [6], researchers used fully connected networks and OCR to generate 3D models from 2D building floor plans. Authors in [10] proposed MapSegNet to segment an indoor map into smaller units including rooms, furniture, corridors etc. They showed that their model achieved better accuracy and lower computation cost than other segmentation methods. In [11], researchers proposed a deep multi-task neural network for floor plan recognition. Using their method, they were able to recognize diverse floor plan elements including walls, rooms, types of rooms, windows etc. They included room-boundary-guided attention mechanisms in their proposed model. In [12], authors implemented a deep neural network to identify junction points from a floor plan. After that, they used integer programming for joining the junctions to generate the walls. In [13], authors recognized walls and openings from the floor plan using heuristics. Then they used the detected walls and doors to generate 3D building models. Also, researchers implemented an improved Hough Transform in [14] to detect wall from floor plan. Their method was based on arc-shaped door hypothesis.

III. METHODS

A. Proposed Model

We customized the Pix2Pix GAN [8] architecture for our experiments. Pix2Pix GAN is a conditional GAN model consists of a generator G and a discriminator D as shown in Fig. 1. In the original paper, authors used “U-Net” [15] as backbone in the architecture for the generator and convolutional “PatchGAN” classifier for the discriminator. In our experiments, we noticed that “U-Net” based generator did not perform well. We then replaced it as the “ResNet” [16] architecture for the generator, which was proposed in Cycle GAN [17]. This is the modification that we have done in the original Pix2pix GAN setting. Fig. 1 represents the overall architecture of the Pix2Pix GAN. Fig. 2 shows the architecture of the generator that was used in our model.

GAN can generate images from noise vectors. Let z be the noise vector and y output image, we can express this mapping as $G: z \rightarrow y$ [8]. On the other hand, Conditional GAN uses input image x and noise vector z to perform the mapping from

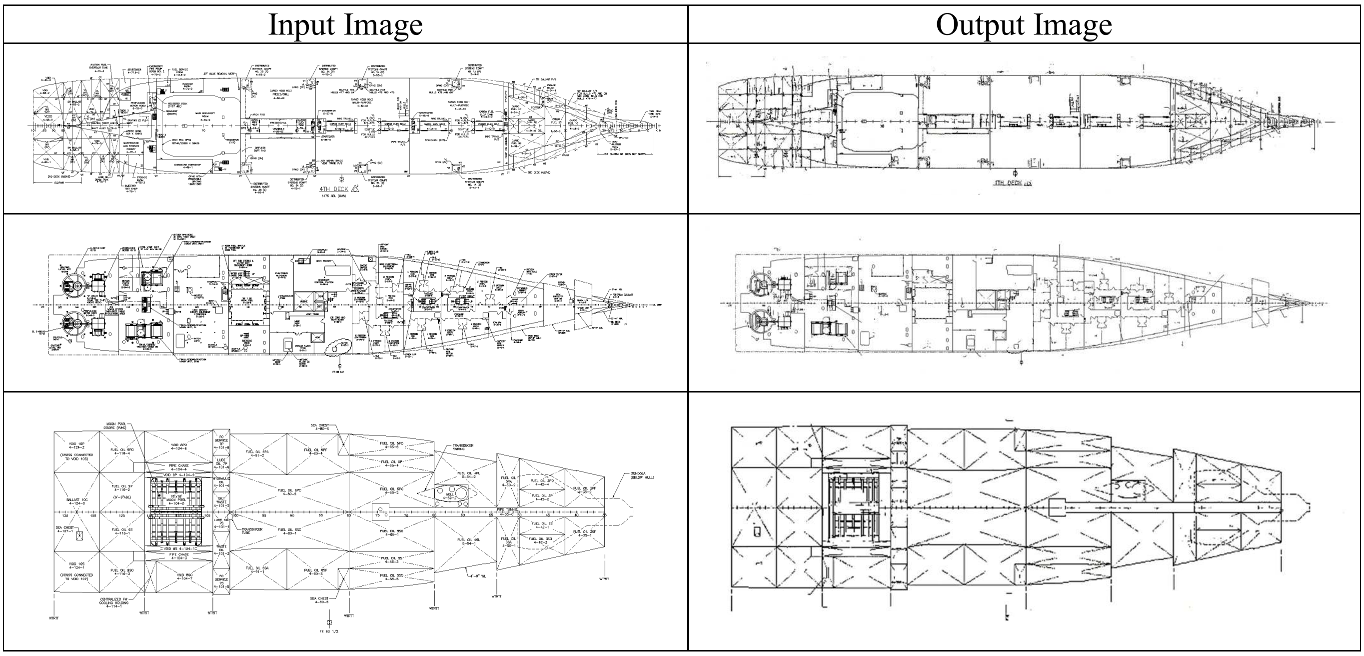


Fig. 5. Test results for ship deck plan segmentation on AutoCAD generated images.

noise z and image x to generate output image y. This mapping can be represented as G: $\{x, z\} \rightarrow y$. Pix2pix GAN is also a conditional GAN model.

The overall objective of the Pix2PIX GAN [8] is given by the following equation where λ is a hyperparameter:

$$G = arg\min_{G}\max_{D} L_{cGAN}(G, D) + \lambda L_{L1}(G) \qquad (1)$$

where

$$L_{cGAN}(G, D) = E_{x,y}[logD(x, y)] + E_{x,z}[\log(1 - D(x, G(x, z)))]$$

$$L_{L1}(G) = E_{x,y,z}\|y - G(x, y)\|_2$$

B. Data Collecdtion for Training

We collected two different datasets of ship deck plans for our experiments. The first dataset had deck plans of ships from AutoCAD software platform [9]. We manually load existing ship floor plan PDF files with many unwanted texts and other annotating drawings in the files into the AutoCAD platform and generated 50 paired clean floor plans as ground truth in which 42 of them were used for training and 8 of them for testing. Fig. 3 shows some example ground truth pairs generated by the AutoCAD platform.

The second dataset consists of MSC engineering documents consisting of scanned images of ship deck plans from traditional PDF files. These old drawings do not have paired ground truth generated by the AutoCAD platform. Fig. 4 shows some samples of this dataset. For this dataset, we manually cleaned the drawings and generated 5 pairs for testing. It is worth noting that the second dataset is more challenging lean than the first one. Moreover, texts and symbols were consistent across all deck plans from AutoCAD files. On the other hand, the MSC engineering documents have variations in symbols and texts styles and it is challenging to clean these images.

C. Training Details

We trained our model using 42 paired images collected from the AutoCAD platform of ship deck plans. We trained our model for 8000 epochs. During training, we randomly cropped 256x256 patches from each image for training and for each iteration we had new patch to train the model.

IV. Experimental results

After training our model, we tested our trained model on both the AutoCAD generated deck plans images (dataset 1) and the manually generated deck plan images (dataset 2). We noticed that we can achieve good performances on images in dataset 1 for segmentation. As the domains of the training and testing images are similar and the results are expected and reasonable. Fig. 5 shows example output results generated by the trained model.

We also applied the trained model to dataset 2 without fine-tuning the model, and results are shown in Fig. 6. It is observed that the model could not perfectly remove texts and labels from the images. The reason behind this result is that there is a domain gap between the training images and testing images for this case. To attack this challenge, we applied pre-processing steps on the testing images including histogram matching and contrast enhancement. The pre-processing steps try to minimize the domain gap between the training images and testing images. After applying pre-processing, we were able to achieve better results as shown in Fig. 7. We have done quantitative comparison as shown in Table 1. We have used three different evaluation matrices for the comparison including dice similarity coefficient, BF score [18], and intersection over union. We can see from the Table that preprocessing had improved the results.

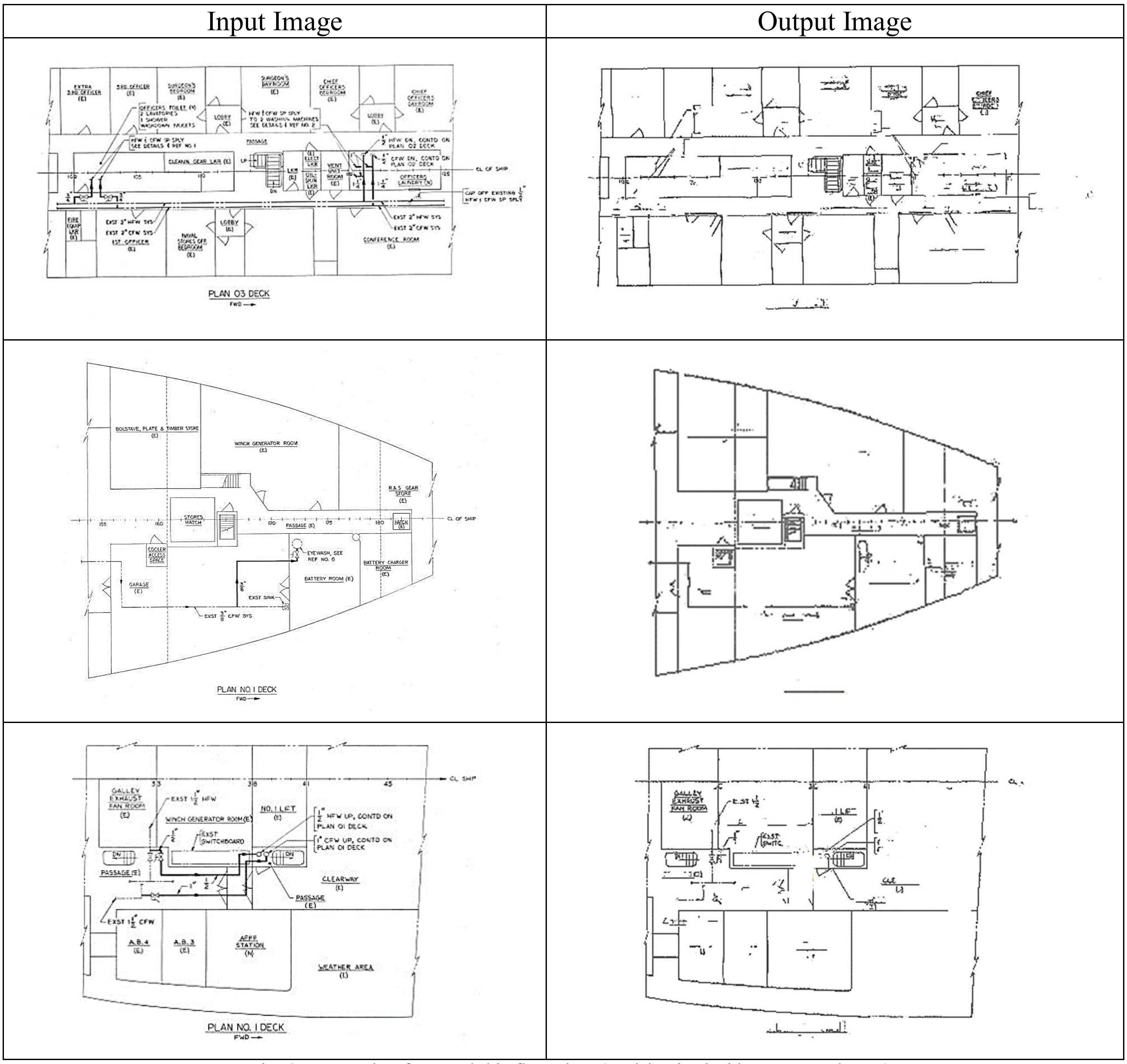


Fig. 6. Test results of scanned ship floor plans (model trained with AutoCAD dataset)

TABLE I. QUANTITATIVE COMPARISON OF OUTPUTS FROM MSC ENGINEERING DOCUMENT FOR SHIP DECK PLAN SEGMENTATION

Evaluation Matrix	***Without preprocessing***	***With preprocessing***
Dice-Similarity Coefficient	0.9821	0.9847
BF (Boundary F1) Score	0.8235	0.9081
Intersection over union	0.9648	0.9701

V. CONCLUSION

In this paper, we have shown that generative adversarial networks can be a good candidate for ship deck plan segmentation. We have performed segmentation on MSC engineering documents to extract ship deck plan. Experiment achieved promising results on testing data having similar domain background as training data. However, the performance degraded on data if there is a domain gap between the testing data and the data used for training the model. Our future work including collecting and manually clean more data from the MSC engineering documents to fine-tune the GAN model and improve the performance of the GAN model on the MSC PDF data.

ACKNOWLEDGMENT

This material is based upon work supported, in whole or in part, by the U.S. Department of Defense through SERC WRT-1045 under contract HQ0034-13-D-0004.

REFERENCES

[1] H. Leon-Garza, H. Hagras, A. Peña-Rios, A. Conway and G. Owusu, "A Fuzzy Rule-based System using a Patch-based Approach for Semantic Segmentation in Floor Plans," 2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2021, pp. 1-6, doi: 10.1109/FUZZ45933.2021.9494427.

[2] Wen Q, Zhu RG. Automatic Generation of 3D Building Models Based on Line Segment Vectorization. Mathematical Problems in Engineering. 2020 Oct 10;2020.

[3] Xu Z, Yang C, Alheejawi S, Jha N, Mehadi S, Mandal M. Floor Plan Semantic Segmentation Using Deep Learning with Boundary Attention Aggregated Mechanism. In2021 4th International Conference on

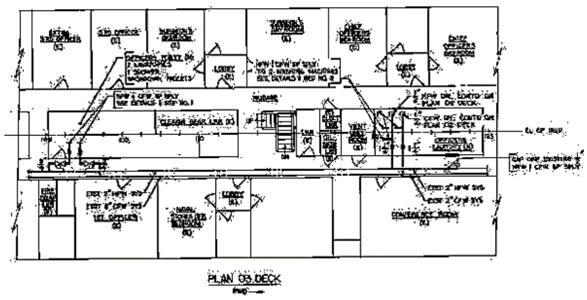
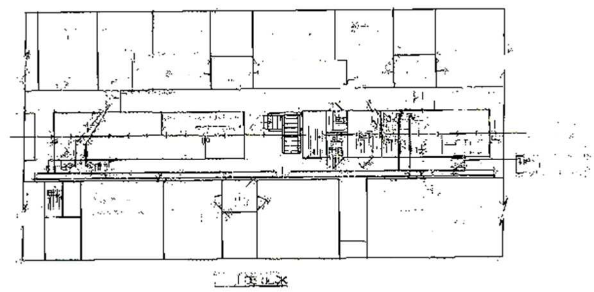
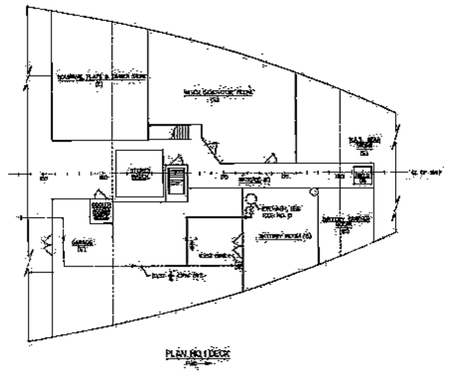
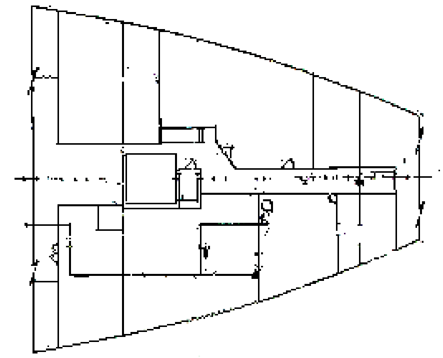
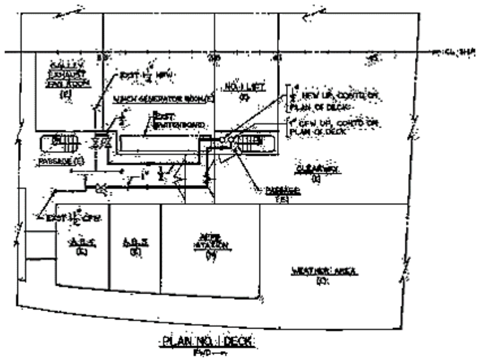
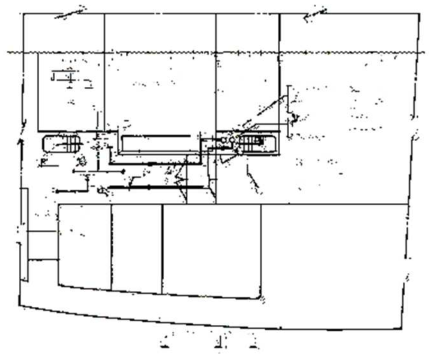
Input Image (After Preprocessing)	Output Image
	
	
	

Fig. 7. Testing results on MSC engineering documents after preprocessing (model trained with AutoCAD dataset).

Artificial Intelligence and Pattern Recognition 2021 Sep 24, pp. 346-353.

[4] Lv X, Zhao S, Yu X, Zhao B. Residential floor plan recognition and reconstruction. InProceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2021, pp. 16717-16726.

[5] Vidanapathirana M, Wu Q, Furukawa Y, Chang AX, Savva M. Plan2Scene: Converting Floorplans to 3D Scenes. InProceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2021, pp. 10733-10742.

[6] Dodge S, Xu J, Stenger B. Parsing floor plan images. In2017 Fifteenth IAPR international conference on machine vision applications (MVA) 2017 May 8, pp. 358-361.

[7] I. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. Advances in neural information processing systems. 2014,27.

[8] Isola P, Zhu JY, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. InProceedings of the IEEE conference on computer vision and pattern recognition 2017, pp. 1125-1134.

[9] 2022. *AutoCAD*. San Rafael, CA: Autodesk.

[10] Foroughi, F., Wang, J., Nemati, A., Chen, Z. and Pei, H.. Mapsegnet: A fully automated model based on the encoder-decoder architecture for indoor map segmentation. *IEEE Access,2021*, *9*, pp.101530-101542.

[11] Zeng, Z., Li, X., Yu, Y.K. and Fu, C.W.. Deep floor plan recognition using a multi-task network with room-boundary-guided attention. In *Proceedings of the IEEE/CVF International Conference on Computer Vision 2019*, pp. 9096-9104.

[12] Liu, C., Wu, J., Kohli, P. and Furukawa, Y.. Raster-to-vector: Revisiting floorplan transformation. In *Proceedings of the IEEE International Conference on Computer Vision* 2017, pp. 2195-2203.

[13] Gimenez, L., Robert, S., Suard, F. and Zreik, K.. Automatic reconstruction of 3D building models from scanned 2D floor plans. *Automation in Construction*,2016, *63*, pp.48-56.

[14] Macé, S., Locteau, H., Valveny, E. and Tabbone, S.. A system to detect rooms in architectural floor plan images. In *Proceedings of the 9th IAPR International Workshop on Document Analysis Systems 2010*, pp. 167-174.

[15] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In MICCAI, 2015.

[16] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR, 2016, pp. 770– 778.

[17] Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. InProceedings of the IEEE international conference on computer vision 2017, pp. 2223-2232.

[18] Csurka G, Larlus D, Perronnin F, Meylan F. What is a good evaluation measure for semantic segmentation?. IEEE PAMI. 2004, 26(1).