TEXT TO IMAGE SYNTHESIS USING GENERATIVE ADVERSERIAL NETWORK

PROJECT REPORT

Submitted in the partial fulfilment of the award of the degree of

Bachelor of Technology

in

Computer Science & Engineering

of

APJ Abdul Kalam Technological University

by

GEORGY JOSE VILAVINAL

MATHEW ALEX

MELBIN MATHEW

VIGNESH HARI

November 2018

Department of Computer Engineering

College of Engineering, Chengannur, Kerala -689121

Phone: (0479) 2454125, 2451424; Fax: (0479) 2451424

COLLEGE OF ENGINEERING, CHENGANNUR
KERALA

Department of Computer Engineering

CERTIFICATE

This is to certify that the seminar entitled

TEXT TO IMAGE SYNTHESIS USING GENERATIVE ADVERSERIAL NETWORK

Submitted by

GEORGY JOSE VILAVINAL

MATHEW ALEX

MELBIN MATHEW

VIGNESH HARI

is a bonafide record of the work done by them.

Co-ordinator **Guide** **Head of the Department**

ACKNOWLEDGEMENT

I am greatly indebted to **God Almighty** for being the guiding light throughout with his abundant grace and blessings that strengthened me to do this endeavour with confidence.

I express my heartfelt gratitude towards **Dr. Jacob Thomas V.**, Principal, College of Engineering Chengannur for extending all the facilities required for doing my seminar. I would also like to thank **Dr. Smitha Dharan**, Head, Department of Computer Engineering, for providing constant support and encouragement.

Now I extend my sincere thanks to my seminar co-ordinators **Mrs. Shiny B**, Assistant Professor in Computer Engineering and **Ms. Archana Vijayan**, Assistant Professor in Computer Engineering for guiding me in my work and providing timely advices and valuable suggestions.

Last but not the least, I extend my heartfelt gratitude to my parents and friends for their support and assistance.

ABSTRACT

Synthesizing high-quality images from text descriptions is a challenging problem in computer vision and has many practical applications. Samples generated by existing text- to-image approaches can roughly reflect the meaning of the given descriptions, but they fail to contain necessary details and vivid object parts. In this paper, we propose Stacked Generative Adversarial Networks (StackGAN) to generate 256x256 photo-realistic images conditioned on text de- scriptions. We decompose the hard problem into more man- ageable sub-problems through a sketch-refinement process. The Stage-I GAN sketches the primitive shape and colors of the object based on the given text description, yield- ing Stage-I low-resolution images. The Stage-II GAN takes Stage-I results and text descriptions as inputs, and gener- ates high-resolution images with photo-realistic details. It is able to rectify defects in Stage-I results and add compelling details with the refinement process. To improve the diversity of the synthesized images and stabilize the training of the conditional-GAN, we introduce a novel Conditioning Augmentation technique that encourages smoothness in the latent conditioning manifold. Extensive experiments and comparisons with state-of-the-arts on benchmark datasets demonstrate that the proposed method achieves significant improvements on generating photo-realistic images condi- tioned on text descriptions.

1 Introduction

2 REFERENCES

[1] J. Altarriba, "Emotion, memory, and bilingualism," in *Foundations of Bilingual Memory,* J. Altarriba and R. Heredia, Eds. New York, NY, USA: Springer, 2014, pp. 185–203.

[2] A. Pavlenko, "Affective processing in bilingual speakers: Disembodied cognition?" *"Int. J. Psychol.,* vol. 47, no. 6, pp. 405–428, 2012.

[3] G. L. Clore and A. Ortony, "The semantics of the affective lexicon," in *Cognitive Perspectives on Emotion and Motivation.* New York, NY, USA: Springer, 1988, pp. 367–397.

[4] E. H. Hovy, "What are sentiment, affect, and emotion? Applying the methodology of Michael Zock to sentiment analysis," in *Language Production, Cognition, and the Lexicon,* N. Gala, R. Rapp, and G. Bel-Enguix, Eds. Cham, Switzerland: Springer, 2015, pp. 13–24.

[5] A.PavlenkoandV.Driagina,"RussianemotionvocabularyinAmerican learners' narratives," *Modern Lang. J.,* vol. 91, no. 2, pp. 213–234, 2007.

[6] C.-C. Huang and L.-W. Ku, "Interest analysis using semantic PageRank and social interaction content," in *Proc. IEEE 13th Int. Conf. Data Mining Workshops,* Dec. 2013, pp. 929–936.

[7] P.-Y.Lu,Y.-Y.Chang,andS.-K.Hsieh,"Causingemotionincollocation: An exploratory data analysis," in *Proc. 25th Conf. Comput. Linguistics Speech Process.,* 2013, pp. 236–249.

[8] B.Pang,L.Lee,andS.Vaithyanathan,"Thumbsup?:Sentiment classification using machine learning techniques," in *Proc. ACL Conf. Empirical Methods Natural Lang. Process.,* 2002, pp. 79–86.

[9] J.-M. Dewaele, "Investigating the psychological and emotional dimensions in instructed language learning: Obstacles and possibilities," *Modern Lang. J.,* vol. 89, no. 3, pp. 367–380, 2005.

[10] J.-M.Dewaele and A.Pavlenko,"Emotionvocabularyininterlanguage," *Lang. Learn.,* vol. 52, no. 2, pp. 263–322, 2002.

[11] T.Kaneko,"Hownon-nativespeakersexpressanger,surprise,anxiety and grief: A corpus-based comparative study," in *Proc. Corpus Linguistics Conf.,* 2003, pp. 384–393.

[12] E. M. Rintell, "That's incredible: Stories of emotion told by second language learners and native speakers," *Developing Communicative Competence in a Second Language.* 1990, pp. 75–94.

[13] J.-M. Dewaele, "On emotions in foreign language learning and use," *Lang. Teacher,* vol. 39, no. 3, pp. 13–15, 2015.

[14] C. R. Graham, A. W. Hamblin, and S. Feldstein, "Recognition of emotion in English voices by speakers of Japanese, Spanish and English," *Int. Rev. Appl. Linguistics Lang. Teach.,* vol. 39, no. 1, pp. 19–37, 2001.

[15] Y. Yeh, H.-C. Liou, and Y.-H. Li, "Online synonym materials and concor- dancing for EFL college writing," *Comput. Assist. Lang. Learn.,* vol. 20, no. 2, pp. 131–152, 2007.

[16] M. Martin, "Advanced vocabulary teaching: The problem of synonyms," *Modern Lang. J.,* vol. 68, no. 2, pp. 130–137, 1984.

[17] D. Liu, "Salience and construal in the use of synonymy: A study of two sets of near-synonymous nouns," *Cognit. Linguistics,* vol. 24, no. 1, pp. 67–113, 2013.

[18] D. Liu and S. Zhong, "L2 vs. L1 use of synonymy: An empirical study of synonym use/acquisition," *Appl. Linguistics,* vol. 37, no. 2, pp. 239–261, Apr. 2016.

[19] W.-F. Chen, M.-H. Chen, and L.-W. Ku, "Embarrassed or awkward? Ranking emotion synonyms for ESL learners' appropriate wording," in *Proc. 10th Workshop Innov. NLP Building Educ. Appl.,* 2015, pp. 144–153.

[20] W.-F. Chen, M.-H. Chen, M.-L. Chen, and L.-W. Ku, "A computer- assistance learning system for emotional wording," *IEEE Trans. Knowl. Data Eng.,* vol. 28, no. 5, pp. 1093–1104, May 2016.