ARTIFICIAL INTELLIGENCE SYNTHESIZED FACE SWAPPING DETECTION MODEL USING UNIFIED DATA SETS

GONG DAFENG

DOCTOR OF PHILOSOPHY

2023

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Faculty of Information and Communication Technology

ARTIFICIAL INTELLIGENCE SYNTHESIZED FACE SWAPPING DETECTION MODEL USING UNIFIED DATA SETS

Gong Dafeng

Doctor of Philosophy

2023

ARTIFICIAL INTELLIGENCE SYNTHESIZED FACE SWAPPING DETECTION MODEL USING UNIFIED DATA SETS

GONG DAFENG

A thesis submitted
in fulfillment of the requirements for the degree of Doctor of Philosophy

Faculty of Information and Communication Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2023

DECLARATION

I declare that this thesis entitled "Artificial Intelligence Synthesized Face Swapping Detection Model Using Unified Data Sets" is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature : 龚大丰

Name : Gong Dafeng

Date : 18/04/2023

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of Doctor of Philosophy.

Signature : ....

Supervisor Name: Dr. Yogan Jaya Kumar

Date : 18/04/2023

DEDICATION

To my beloved father, mother, wife, sisters, brothers, and daughters

ABSTRACT

Today's image generation technology can generate high-quality face images, and it is not easy to recognize the authenticity of the generated images through human eyes. Due to the rise of image generation technology based on deep learning, software related to image generation is used widely, including some popular face-swapping software. If misused, it will directly affect forensics and security-related industries. As an essential branch of computer security, image forensics technology also needs to be improved with the development of image forgery technology. This study aims to improve deepfake detection, a face-swapping forgery, by absorbing the advantages of deep learning technologies. In order to solve the problem of poor detection performance on cross data sets, this study generates unified data sets from multiple sources using spatial enhancement technology to obtain approximately four million images, 36 times the size of the original data set, and was proved effective with traditional feature methods. Taking the advantages of ResNet and Inception networks, DeepfakeNet architecture composed of 32 parallel branches and 20 network layers is proposed as the deepfake detection model with FLOPs of 2.05×10^9 and parameters of 10.87×10^6. To further improve the proposed DeepfakeNet model, a univariate method is used to obtain the ideal model values of hyperparameters, including batch size, epochs, dropout, learning rate, and sample ratio. Accuracy of 98.69%, loss value of 3.42% and AUC of 0.96 are achieved. The evidence of this study shows that the proposed DeepfakeNet has significantly improved over the mainstream methods in terms of loss value, accuracy, AUC, FLOPs, and parameters.

MODEL PENGESANAN PERTUKARAN MUKA MELALUI KECERDASAN BUATAN MENGGUNAKAN SET DATA BERSATU

ABSTRAK

Teknologi penjanaan imej hari ini boleh menjana imej muka berkualiti tinggi, dan bukan mudah untuk mengenali ketulenan imej yang dijana melalui mata manusia. Disebabkan oleh peningkatan teknologi penjanaan imej berdasarkan pembelajaran mendalam, perisian yang berkaitan dengan penjanaan imej digunakan secara meluas, termasuk beberapa perisian pertukaran muka yang popular. Jika disalahgunakan, ia akan menjejaskan industri forensik dan berkaitan keselamatan secara langsung. Sebagai cabang penting dalam keselamatan komputer, teknologi forensik imej juga perlu dipertingkatkan dengan pembangunan teknologi pemalsuan imej. Kajian ini bertujuan untuk meningkatkan pengesanan deepfake, pemalsuan pertukaran muka, dengan menyerap kelebihan teknologi pembelajaran mendalam. Untuk menyelesaikan masalah prestasi pengesanan yang lemah pada set data silang, kajian ini menjana set data bersatu daripada pelbagai sumber menggunakan teknologi peningkatan spatial untuk mendapatkan kira-kira empat juta imej, 36 kali ganda saiz set data asal, dan telah terbukti berkesan dengan kaedah ciri tradisional. Mengambil kelebihan rangkaian ResNet dan Inception, seni bina DeepfakeNet yang terdiri daripada 32 cawangan selari dan 20 lapisan rangkaian dicadangkan sebagai model pengesanan deepfake dengan FLOP 2.05×10^9 dan parameter 10.87×10^6. Untuk menambah baik lagi model DeepfakeNet yang dicadangkan, satu univariate kaedah digunakan untuk mendapatkan nilai model ideal hiperparameter, termasuk saiz kelompok, zaman, keciciran, kadar pembelajaran, dan nisbah sampel. Ketepatan 98.69%, nilai kehilangan 3.42% dan AUC sebanyak 0.96 dicapai. Bukti kajian ini menunjukkan bahawa DeepfakeNet yang dicadangkan telah bertambah baik dengan ketara berbanding kaedah arus perdana dari segi nilai kehilangan, ketepatan, AUC, FLOP dan parameter.

ACKNOWLEDGEMENTS

Firstly, I would like to express my gratitude to my supervisors, Dr. Yogan Jaya Kumar and Professor Ts. Dr. Goh Ong Sing, for their continuous support and availability in the development of my research study, especially because both of them have never been too busy to keep an eye on my progress in spite of their numerous obligations. They have greatly helped me in a lot of ways throughout this study. The most important lesson I learnt from them is that whatever you work on, work with dedication and sincerity. Again, I owe them my deepest thanks.

In addition, I am grateful to my colleagues at my research lab, especially those in a deep learning research group, for the useful discussions, comments, and knowledge sharing. Thanks also to my friends here at UTeM, who have in one way or another, helped, encouraged, and motivated me during the progression of my study.

I am also grateful to all staff of Faculty of Information and Communication Technology and other departments, UTeM, for their kind cooperation during my study and stay here.

Last but not least, I like to thank my parents, wife, sisters, brothers and daughters for their support, guidance, love and prayers. Thank you for all the encouragement you have given me. Once again, thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

AI	-	Artificial Intelligence
BGD	-	Batch Gradient Descent
BN	-	Batch Normalization
CFFN	-	Common Fake Feature Network
CNN	-	Convolutional Neural Networks
DCT	-	Discrete Cosine Transform
DFAE	-	Deepfake Autoencoder
DFD	-	Deepfake Detection
DFDC	-	Deepfake Detection Challenge
DL	-	Deep Learning
DWT	-	Discrete Wavelet Transform
ELA	-	Error Level Analysis
EXIF	-	Exchangeable Image File Format
FC	-	Full Convolution
FCN	-	Full Convolution Neural Network
FF	-	FaceForensics
FFmpeg	-	Fast Forward Moving Picture Expert Group
FLOPs	-	Floating Point Operations
FN	-	False Negative
FP	-	False Positive
GAN	-	General Adverse Network
GPU	-	Graphic Processing Unit
IPM	-	Integral Probability Metrics
LAE	-	Locality-aware AutoEncoder

LBP	-	Local Binary Pattern
LR	-	Learning Rate
LSTM	-	Long Short-Term Memory
MBGD	-	Mini-Batch Gradient Descent
ML	-	Machine Learning
MLP	-	Multilayer Perceptron
PDR	-	Predictive, Descriptive, Relevant
PRNU	-	Photo Response Non-Uniformity
RAM	-	Random Access Memory
RF	-	Random Forest
RGB	-	Red-Green-Blue
RNN	-	Recurrent Neural Network
SGD	-	Stochastic Gradient Descent
SVD	-	Singular Value Decomposition
SVM	-	Support Vector Machines
TN	-	True Negative
TP	-	True Positive
VDM	-	Variable Divergence Minimization
VGG	-	Visual Geometry Group

LIST OF SYMBOLS

$Tanh$	-	Tanh activation function
$ReLU$	-	ReLU activation function

LIST OF PUBLICATIONS

1. Gong, D., Sing, G. O., Kumar, Y. J., Ye, Z. and W, Chi., 2020. Deepfake Forensics, an AI-synthesized Detection with Deep Convolutional Generative Adversarial Networks. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(3), pp. 2861–2870.

2. Gong, D., Sing, G. O., Kumar, Y. J., Ye, Z. and W, Chi., 2021. DeepfakeNet, an Efficient Deepfake Detection Method. *International Journal of Advanced Computer Science and Applications*, 12(6), pp. 201–207. (SCOPUS indexed)

3. Gong, D., Kumar, Y. J., Sing, G. O., Choo, Y. H., Ye, Z. and W, Chi., 2022. An Improved Deepfake Detection Method Based on CNNs. *Journal of Theoretical and Applied Information Technology*, 100(18), pp. 5684-5691. (SCOPUS indexed)

4. Jin Z., Liu L., Gong D., Li L., 2021, Target Recognition of Industrial Robots Using Machine Vision in 5G Environment. *Frontiers in Neurorobotics*, 15, pp. 1-9. (SCOPUS indexed, SCI indexed)

5. Gong D., Cai S., Chi W., 2022. Research on the feedback teaching of AI customized development. *Scientific Journal of Intelligent Systems Research*, 4(7), pp. 700-708.

6. Chi W., Choo, Y. H., Sing, G. O., Gong D., 2022. Lung Disease Diagnosis based on Transfer Learning. *Journal of Artificial Intelligence Practice*. 5(1), pp. 98-104.

7. Gong, D., Kumar, Y. J., Sing, G. O., Ye, Z. and W, Chi., 2021. Data Enhancement Technology on Deepfake Dataset. in *3rd International Conference on Intelligent and*

Interactive Computing 2021, pp. 1-2.

8. Ye, Z., Kumar, Y. J., Sing, G. O. and Gong, D., 2021. Research on Upgrading of Traditional Industry Driven by Intelligent Manufacturing. in *3rd International Conference on Intelligent and Interactive Computing 2021*, pp. 85.