Deepfakes, a Threat to Society

Mrs. Usha P. Kosarkar[1], Dr. Shilpa R. Gedam[2], Dr. Gopal Sakarkar[3]
[1]Department of Computer Science, G.H. Raisoni University, Saikheda, Maharashtra, India
[2]Department of Computer Science, SSESA's Science College, Nagpur, Maharashtra, India
[3]Department of Computer Science, GHRCE, Nagpur, Maharashtra, India

ABSTRACT

Nowadays, people faced a problem of face swapping images and forged videos, widely known as the Deepfakes. These kind of images and videos are being circulated on social media , freely causing problem peoples privacy. Some deepfake images are very hard to distinguish from original ones and cannot be identified by human eye. This concept of fabrication and manipulation of digital videos and images are not new. This paper discusses about the fact of face swapping algorithms , their impact on the media, a review of deepfake and its development over the years. Conclusion of this paper offers recommendations based on the analysis.

Keywords: Deepfakes, Generative Adversarial Networks(GANs), deepfake threats Machine Learning.

I. INTRODUCTION

In the recent times, multimedia is used as a tool for alteration and manipulation. This altered and manipulated multimedia is freely circulated on the social media platforms without any hesitation [7]. The concept of deepfake was invented in 2014 by Ian Good fellow. Deep fakes are produced using Artificial Intelligent (AI) applications and Machine Learning that merge, combine, replace, and superimpose images and video clips to make fake videos that appear as if they are original ones [2].

Recently Social Networking sites like FaceBook and Instagram have announce their policy in January 2020 regarding banning of deepfake videos. There are many examples of Superimposing someone's face with someone else's. Specially faces of celebrities are used for this purpose to tarnish their image in society. Like in a photo the U.S.A. president Lincoln's head was swapped with politician John Calhoun's was produced in mid-19th century.

A study recently published in Cognitive Research[1] tried to measure people's ability to recognize whether a photo has been manipulated or not. The study showed that only 62% to 66% of the photos were correctly classified as original or manipulated ones. In a similar study published by Harisha et al [2] that only 58% of the images were correctly classified as original and only 46% of the images were identified as manipulated ones. The threat represented by widespread image forgery has stimulated intense research in multimedia forensics. Because of this we need an automatic algorithms for better detection of original and manipulated images than humans. For example, in the first IEEE Image Forensics Challenge, detection accuracies beyond 90% were obtained by means of a machine learning approach with a properly trained classifier [3].

However, when the idea of neural network became popular, people began to use this technology in their everyday life. Subsequently, these techniques have been used by artists, pranksters and many others to create a collection of audio and video files depicting high-profile leaders, like Donald Trump Vladimir Putin and Barack Obama, saying things they never did. The trend has inevitably instilled fears within the national security community. Same technology was also used to create forged pornographic content, which was a threat to society.

After examining the technical literature available on deepfakes in order to assess the threat they pose, the paper draws two conclusions. Firstly, generating crude deepfakes for malicious use will become easier with time as the technology commodifies. At the same time though, the current situation of deepfake detection suggests that we can largely keep these fakes at bay. Secondly, the greater threat will come from tailored deepfakes produced by technically sophisticated actors.

People in general have a broad, vague fear that synthetic media will eventually destroy our ability to identify the real from the fake. According to one New York Times op-ed writer in 2019: “Deepfakes Are Coming. We Can No Longer Believe What We See.”

Face-swapping involves the automatic replacement of a face in a video or image with another face where the identity of the person in the video changes. This original face-swapping method can be dated back to a Reddit user post in 2017 [10]. Faceswap-GAN is a popular face swap method [8]. Based on the original deepfakes method, Faceswap-GAN adds antagonistic and perceptual loss to the result of the automatic coding system. Adding counter losses improves the reconstruction quality of the generated image. The addition of perceptual loss improves eye orientation and aligns the face of the generated image with the input image. This method is an optimized version of the original deepfakes approach [1].

II. TECHNOLOGY USED

Most of the deepfakes are created with powerful graphics cards or with better computing power. This reduces the time interval from days and weeks to hours. But it takes expertise, too,not least to the touch up completed videos to scale back flicker and other visual defects. That said, many tools are now available to assist people make deepfakes. Several companies will make them for you (deepfakesweb.com) and do all the processing within the cloud. There’s even a mobile app, Zao which lets users overlap their faces to a long list of TV and movie characters on which the system has trained. [2]

Similarly, government discourse on these issues have been shaped by broad concerns. The chairman of a congressional committee, after listing some possible malicious uses at a hearing on the matter, acknowledged that it is not too difficult to imagine even more horrifying scenarios that would leave the government, the media, and the public struggling to identify real from fake in future.

Since deep neural networks have been widely used in various recognition tasks, we can also adopt a deep neural network to detect fake images generated by the GANs. Recently, the deep learning-based approached for fake image detection using supervised learning has been studied. In other words, fake image detection has been treated as a binary classification problem (i.e., fake or real image). For instance, the convolution neural network (CNN) network was used to develop the fake image detector [9,10]. In [11], the performance of the fake face image detection was further improved by adopting the most advanced CNN–Xception network [12].

In [13], a manipulated face detection algorithm was proposed based on a hybrid ensemble learning approach. However, none of these studies has investigated the fully generated image, but instead, they have been focused only on partial manipulation of face images; thus, they cannot be used to detect the fully generated fake images. Many GANs have been proposed in recent years. Some of the recently proposed GANs [1–3,14–18] have been used to produce photo-realistic images. To develop a fake image detector, it is necessary to collect all of the GAN's images as the training set for deep neural networks to achieve the promising performance. However, it is difficult and very time-consuming to collect the training samples generated by all the GANs. In addition, such a supervised learning strategy [9–11] tends to learn the discriminative features of fake images generated by all the GANs, and as a result, the learned (trained) detector may not have a good generalization ability. In other words, the learned detector will be unable to recognize the fake images generated by the GANs that were not included in the detect or training process. To meet the current requirement for the GANs-based generator of fake image detection, WorkuMuluyeWubet proposed a modified network structure, including a pairwise learning approach, called the common fake feature network (CFFN)[4].

There is also positive use of deepfakes such as creating voices of those who have lost theirs or updating episodes of movies without reshooting them [14]. However, the number of malicious uses of deepfakes largely dominates that of the positive ones. The development of advanced deep networks and the availability of large amount of data have made the forged images and videos almost indistinguishable to humans and even to sophisticated computer algorithms. The process of creating those manipulated images and videos is also much simpler today as it needs as little as an identity photo or a short video of a target individual. Less and less effort is required to produce a stunningly convincing tempered footage. Recent advances can even create a deepfake with just a still image [5].

Additionally, several techniques to detect videos containing facial manipulations have been presented. While some of these methods focus on detecting videos containing only DeepFake manipulations, others are designed to be agnostic to the technique used to perform the facial manipulation. The work presented in [30, 31] use a temporal-aware pipeline composed by a Convolutional Neural Network (CNN) and a Recurrent Neural Network (RNN) to detect DeepFake videos. Current DeepFake videos are created by splicing synthesized face regions onto the original video frames. This splicing operation can leave artifacts that can later be detected when estimating the 3D head pose. The authors of [32] exploit this fact and use the difference between the head pose estimated with the full set of facial landmarks and a subset of them to separate DeepFake videos from real videos. This method provided competitive results on the UADFV [33] database. The same authors proposed a method [34] to detect DeepFake videos by analyzing the face warping artifacts. The authors of [20] detect manipulated videos generated by the DeepFake and Face2Face techniques with a shallow neural network that acts on mesoscopic features extracted from the video frames to distinguish manipulated videos from real ones. However, the results presented in [21] demonstrated that in a supervised setting, several deep network based models [35, 36, 37] outperform the ones based on shallow networks when detecting fake videos generated with DeepFake, Face2Face, FaceSwap, and Neural Texture[7].

III. THREE SIGNIFICANT FACTORS

Three key factors that will shape the use of technology in future are - the compelling feature of ML-driven faux media, the operational requirements of using the technology, and the risks of identification and detection raised by using deepfakes.

A. Advantage: Compelling Feature

Deepfakes give a unique opportunity to the online campaigner in order to create deceitful content. ML-based duping can generate strikingly realistic portrayal of individuals Center for Security and Emerging Technology 3 and situations. Especially, deepfakes can replicate subtle and minute details like convincing facial tics or realistic shadows for a fake object pasted into an image identifying fake images becomes extremely difficult due to these details. Nonetheless, these fakes are enough to bring confusion and suspicion about the targeted individual or situation. Numerous examples of crudely produced fakes which are widely circulated and perceived as real can be abundantly found on the internet. Consider the 2019 video of Speaker of the House Nancy Pelosi that spread extensively through social media, purporting to show Pelosi either drunk or suffering from some kind of mental deterioration. No ML was used in this case. The video was produced simply by slowing down a real video of Pelosi speaking at an event.

The need to achieve visual realistic fakes is clearly not required for successful hoaxes by malicious actors. The more important factor in the success of a hoax image is clearly based on "motivated reasoning" i.e. the tendency to accept information confirming pre-existing prejudices. Hence, deepfakes is in fact an unappealing method for spreading false descriptions, especially when you weigh the costs and risks involved in using this technology.

B. Expenditure: Operational Requirements

Maligning disinformation campaigns will have to bear certain operational costs in order to adopt ML. Invariably, creating high-performance AI systems needs access to a sufficient training data (enabling a machine to learn how to accomplish a given task) and computational power (the hardware needed to execute the training process). With respect to the depiction of the content of the deepfake, inevitable and high expenditure will be incurred for acquiring the training data, structuring it properly, and running the training process. At the same time, it is becoming more and more convenient to work with software platforms with integrated deepfake technologies. For example, no technical expertise is required on the users' end in order to work on easy-to-use, ML-driven software that can remove one face from an image or video and insert another, also commonly known as "face swap".

C. Perils: Algorithmic Detection

Avoiding public exposure is preferred by influence operations. An influence campaign can be "deplatformed" by social media companies by discovering, deleting accounts and hampering access to users by malicious actors. By using deepfakes, in fact online influence operations may increase their risk of exposure. Hence, deepfakes may contain a kind of "fingerprint," which allows investigators to link together all media from a given disinformation campaign. Investigators, in turn, can trace the campaign to a specific source and alert the public. The distribution of their content through intermediaries, such as Facebook, Twitter, and YouTube. As fears over deepfakes have escalated, these platforms have created new policies prohibiting the use of certain kinds of synthetic media. These policies will use detection algorithms for enforcement, given the massive scale of content uploaded and shared on social media. By choosing to distribute deepfakes, influence operations run the risk of their messaging being quickly taken down or flagged as suspicious on these platforms. These increased risks of exposure and detection may make deepfakes a less attractive means of spreading false narratives than existing methods. Manually copying content from many sources and editing media as needed may avoid the consistent "fingerprints" left by ML models.

The adoption of deepfakes for disinformation purposes will therefore depend on more than the costs of producing this content and its likely impact on the target audience. It will also depend on the speed of improvement in deepfake detection and the adoption of detection technologies by online platforms, governments, and everyday users.

IV. ANALYSIS

Three key factors determine whether and how online influence operations will use deepfakes:

What can be depicted in a deepfake:

1) Disinformation actors will adopt ML only if it creates synthetic media likely to shape public perceptions or cast doubt in the minds of a target audience.
2) The computational, human, and data requirements of generating deepfakes.: High costs of production will make deepfakes less attractive relative to manual methods, while low production costs will make them more attractive.
3) The effectiveness of detection systems: The ability to detect deepfakes at low cost makes ML less attractive to disinformation actors, while ineffectual or high costs to detection make it more attractive.

V. THE STATE OF PLAY

The state of the ML field will define the persuasive capacity, operational requirements, and detection risks of deepfakes.

1. Deepfake Creation

Deepfakes must meet two criteria in order for online influence campaigns to use them. First, the operational costs of producing a deepfake—buying hardware, acquiring data, and hiring expert engineers—must not be overly onerous. Second, deep generative models must be able to successfully produce the faked media an influence campaign seeks to distribute.

2. How to Build a Deepfake

Deepfakes are one specific application of ML, a field focused on the development of algorithms that improve as they process data. This processing results in a trained "model," a piece of software that ideally accomplishes the desired task. The first step is to bring together a training dataset of both tagged photographs of faces and photographs containing no faces.

The ML algorithm learns from the provided examples to associate the images containing a face with the tag "face" and images without faces with the tag "no face." This model gains a limited "understanding" of what a face looks like through the training process. This level of "understanding" is referred to in the field as a representation. Representations are at the core of how ML creates synthetic media.

Specifically, engineers create faked media using a generative model—a class of models that can produce novel data similar to that used to train the system in the first place. This representation can then produce new images of faces that have never existed. The imitations produced by deep generative models are the "deepfakes" sparking public concern. This is an extremely active area of research: numerous models have been proposed in recent years that adopt different approaches with varying strengths and weaknesses. Some of the most prominent examples focus on the generation of images, including Glow (2018), PixelCNN (2016), NADE (2016), and DRAW (2015)[22-28].

One technique a major source of the "deepfakes" most widely circulated beyond the research community is known as generative adversarial networks, or GANs. . Therefore, a GAN with a discriminator trained on images of faces would produce a generative network that can create novel, synthetic images of faces.

3. Costs and Capabilities

Examining the technical literature on generative models helps determine the resources required to produce a high-quality deepfake, and the range of different kinds of faked media that can be generated.With models trained 10 Center for Security and Emerging Technology on faces, malicious actors might seek to produce believable profile photos for fake accounts on social media platforms or to create a false narrative around a made-up individual. One widely-cited paper from 2017 illustrates that state-of-the-art GANs can produce realistic, synthetic face images up to a 1024 x 1024 pixel resolution.

A disinformation campaign unwilling to deal with the cost and complexity of creating a deepfake from scratch could obtain a pre-trained model created by someone else. Increasingly, pre-trained models are being open-sourced or embedded in software for use by laypeople. Today, the basic technology for creating fake swaps is now freely available in open-source software repositories online.48 Freely or cheaply available generative models for creating a range of different fakes will likely become the norm as the knowledge to create deepfakes grows more widespread.

Online influence campaigns will not make the decision to use deepfakes in a vacuum. . Disinformation campaigns will avoid easily detectable deepfakes in favor of ones harder to identify.

VI. CONCLUSION

A dramatic demonstration in the lab often reveals little about how a technology will be used in the real world. Deepfakes are no exception. While the use of ML to produce sharp, high-fidelity synthetic media is an impressive technical feat, the incentives of malicious actors will shape the ultimate threat the technology poses. Policymakers and national security researchers should avoid giving in to hype, but rather take precautions when sensible.

Deepfakes are not magic: ML is not yet so advanced that it can effortlessly conjure up fake scenes indistinguishable from reality. There is a real cost in using ML. Training data, computational power, and technical expertise must all be assembled to use it effectively. Limitations in the methodology constrain what fakes can be made, and how quickly they can be generated. Moreover, constantly evolving detection methods can make synthetic media easier to identify "in the wild." These real, somewhat humdrum considerations provide crucial hints toward how a disinformation campaign is likely to use this technology to manipulate public discourse.

While commodification will make deepfakes ever easier to produce, off-the-shelf technology for producing synthetic media will also become easier to detect and filter automatically. This limits the impact of this technology on mainstream platforms and narrows their scope to less monitored areas of the web. The greater threat is likely from a sophisticated disinformation effort that tailors ML models for particular purposes. Moderately well-resourced disinformation efforts can afford custom generative models that produce cuttingedge deepfakes, but even in these cases, malicious actors are Conclusion A 30 Centre for Security and Emerging Technology constrained. The strategic dynamics of detection, the demands of training time, and accessibility of data all conspire to make some operational uses of deepfakes likelier than not.

VII. REFERENCES

[1] . Deng Pan, Lixian Sun, Rui Wang, Xingjian Zhang, Richard O. Sinnott, Deepfake Detection through Deep Learning, 2020 IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (BDCAT), School of Computing and Information Systems The University of Melbourne, Melbourne, Australia

[2] . Hrisha Yagnik1, Akshit Kurani2, Prakruti Joshi3, A Brief Study on Deepfakes, International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 07 Issue: 12 | Dec 2020 www.irjet.net p-ISSN: 2395-0072 , Department of Computer Engineering, Indus University, Ahmedabad,India.

[3] . WorkuMuluyeWubet, The Deepfake Challenges and Deepfake Video Detection. International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-6, April 2021

[4] . Peisong He1, Haoliang Li2, Hongxia Wang1, Detection of fakeimages via the ensemble of deep representations From multicolor spaces. College of Cybersecurity, Sichuan University, Chengdu, China 2School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

[5] . Artem A. Maksutov1, Viacheslav O. Morozov, Aleksander A. Lavrenov, Alexander S. Smirnov, Methods of Deepfake Detection Based on Machine Learning, Department of Computer Systems and Technology National Research Nuclear University “MEPhI” Moscow, Russian Federation

[6] . ThanhThi Nguyen, Cuong M. Nguyen, Dung Tien Nguyen, DucThanh Nguyen, SaeidNahavandi, Fellow, IEEE. Deep Learning for Deepfakes Creation and Detection: A Survey, arXiv:1909.11573v2 [cs.CV] 28Jul 2020

[7] . DigvijayYadav,SakinaSalmani, Deepfake: A Survey on Facial Forgery Technique Using Generative Adversarial Network, Dept of Master of Computer Applications Sardar Patel Institute of Technology Sardar Patel Institute of Technology Mumbai, India.

[8] . Teng Zhang, Lirui Deng,Liang Zhang, Xianglei Dang: 2020 IEEE 3rd International Conference on Computer and Communication Engineering Technology. Department of Computer Science and Technology Tsinghua University Beijing, China.Technology Research Division CNCERT/CC Beijing, China

[9] . Daniel Mas Montserrat, HanxiangHao, S. K. Yarlagadda, SriramBaireddy, Ruiting Shao J´anosHorv´ath, Emily Bartusiak, Justin Yang, David G¨uera, Fengqing Zhu, Edward J. Delp: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Video and Image Processing Laboratory (VIPER) School of Electrical Engineering Purdue University West Lafayette, Indiana, USA.

[10] . Francesco Marra, Diego Gragnaniello, DavideCozzolino, Luisa Verdoliva: Dep. of Electrical Engineering and Information Technology University Federico II of Naples Naples, Italy”: 2018 IEEE Conference on Multimedia Information Processing and Retrieval

[11] . SiweiLyu: Deepfake detection:Current challenges and next steps: ComputerScienceDepartmentUniversityatAlbany,StateUniversityofNewYork: 978-1-7281-1485-9/20/$31.00 c 2020IEEE

[12] . Badhrinarayan Malolan, Ankit Parekh, FarukKazi: Explainable Deep-Fake Detection Using Visual Interpretability Methods: Centre of Excellence (CoE) in Complex and Non-linear Dynamical Systems (CNDS), VeermataJijabai Technological Institute Mumbai, India: 2020 3rd International Conference on Information and Computer technologies (ICICT)

[13] . Md Rafiqul Islam ,Shaowu Liu, Xianzhi Wang,Guandong Xu: Deep learning for misinformation detection on online social networks: a survey and new perspectives: Received: 28 March 2020 / Revised: 11 September 2020 / Accepted: 12 September 2020 / Published online: 29 September 2020 © Springer-Verlag GmbH Austria, part of Springer Nature 2020

[14] . Nikita S. Ivanov, Anton V. Arzhskov, Vitaliy G. Ivanenko: Combining Deep Learning and Super-Resolution Algorithms for Deep Fake Detection: Department of Computer Systems and Technologies National Research Nuclear University MEPhI (Moscow Engineering Physics Institute) Moscow, Russian Federation: 978-1-7281-5761-0/20/$31.00 ©2020 IEEE

[15] . Mohammed A. Younus,Taha M. Hasan: Abbreviated View of Deepfake Videos Detection Techniques: Department of Computer Science College of Science, University of DiyalaDiyala , Iraq: 6th international engineering conference" Sustainable Technology and Development” , (IEC-2020),

[16] . Md. ShohelRana,Andrew H. Sung: DeepfakeStack: A Deep Ensemble-based Learning: Computing Sciences and Computer Engineering The University of Southern Mississippi Hattiesburg, MS 39406, United States: 2020 7th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2020 6th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom)

[17] . Weiguo Zhang and Chenggang Zhao:: Exposing Face-Swap Images Based on Deep Learning and ELA Detection : College of Computer Science and Technology, Xi’an University of Science and Technology Published: 17 November 2019

[18] . Dongyue Chen , Qiusheng Chen , Jianjun Wu , Xiaosheng Yu , and Tong Jia: Face Swapping: Realistic Image Synthesis Based on Facial Landmarks Alignment: College of Information Science and Engineering, Northeastern University, China 2 Faculty of Robot Science and Engineering, Northeastern University, ChinaL: Hindawi Mathematical Problems in Engineering Volume 2019, Article ID 8902701, 11 pages https://doi.org/10.1155/2019/8902701

[19] . D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen,“Mesonet: a compact facial video forgery detection network,” Proceedings of the IEEE International Workshop on Information Forensics and Security, pp. 1–7, December2018, Hong Kong. 2

[20] . A. Rossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, “Faceforensics++: Learning to detect manipulated facial images,” Proceedings of the IEEE International Conference on Computer Vision, pp. 1–11, October 2019, Seoul, South Korea. 2, 6.

[21] . Z. Hui, J. Li, X. Wang, and X. Gao, “Image fine-grained inpainting,” arXiv preprint arXiv:2002.02609, 2020.

[22] . H. Le and D. Samaras, “Shadow removal via shadow image decomposition,” Proceedings of the IEEE International Conference on Computer Vision, pp. 8578–8587, October 2019,Seoul, South Korea.

[23] . A. Brock, J. Donahue, and K. Simonyan, “Large scale GAN training for high fidelity natural image synthesis,” arXiv preprint arXiv:1809.11096, 2018.

[24] . "DeepFakes," https://github.com/deepfakes/faceswap. J. Thies, M. Zollhofer, and M. Nießner, "Deferred neural ̈rendering: Image synthesis using neural textures," ACM Transactions on Graphics, vol. 38, no. 4, pp. 1–12, July 2019.

[25] . M. Kowalski, "Faceswap," https://github.com/ MarekKowalski/FaceSwap/. J. Thies, M. Zollhofer, M. Stamminger, C. Theobalt, and M. Nießner, "Face2face: Real-time face capture and reen actment of rgb videos," in Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, LasVegas, NV, June 2016, pp. 2387–2395.

[26] . P. Perez, M. Gangnet, and A. Blake, "Poisson image editing," Proceedings of the ACM Special Interest Group on Computer GRAPHics and Interactive Techniques, pp. 313–318, July 2003, San Diego, California. 2

[27] . D. Guera and E. J. Delp, "Deepfake video detection using ̈recurrent neural networks," Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance, pp. 1–6, November 2018, Auckland, New Zealand. 2, 6

[28] . E. Sabir, J. Cheng, A. Jaiswal, W. AbdAlmageed, I. Masi, and P. Natarajan, "Recurrent convolutional strategies for face manipulation detection in videos," Interfaces (GUI), vol. 3,pp. 1, 2019. 22858 Authorized licensed use limited to: Carleton University. Downloaded on August 08,2020 at 06:41:05 UTC from IEEE Xplore. Restrictions apply.

[29] . X. Yang, Y. Li, and S. Lyu, "Exposing deep fakes usininconsistent head poses,"Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 8261–8265, May 2019, Brighton, United Kingdom.

[30] . Y. Li, M.-C. Chang, and S. Lyu, "In ictu oculi: Exposing ai created fake videos by detecting eye blinking," Proceeding IEEE International Workshop on Information Forensics and Security, pp. 1–7, 2018, Hong Kong.

[31] . Y. Li and S. Lyu, "Exposing deepfake videos by detecting face warping artifacts," arXiv preprint arXiv:1811.00656, 2018. 2

[32] . G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, pp. 4700–4708, July 2017, Honolulu, HI. 2

[33] . C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna,"Rethinking the inception architecture for computer vision," Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, pp. 2818–2826, June 2016, Las Vegas, NV. 2

[34] . F. Chollet, "Xception: Deep learning with depthwise separable convolutions," Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, pp. 1251–1258, July 2017, Honolulu, HI. 2, 6

[35] . J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei-Fei, "Imagenet: A large-scale hierarchical image database," pp. 248–255, August 2009, Miami, FL. 3, 4

[36] . M. Huh, P. Agrawal, and A. A. Efros, "What makes imagenet good for transfer learning?," arXiv preprint arXiv:1608.08614, 2016. 3

[37] . K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," IEEE Signal Processing Letters, vol. 23, April 2016. 3