Quantum Blend: the Next Frontier of Artificial Intelligence(AI)-Driven Hyper Realism

Adil Husain Rather, Inam Ul Haq, Malik Muzamil Ishaq, Navjot Singh, Vikash Yadav and Karan Sarawagi

July 4, 2024

Quantum Blend: The Next Frontier of Artificial Intelligence(AI)-Driven Hyper Realism

1st Adil Husain Rather
Dept of Computer Science Engineering
Chandigarh University
Punjab, India
adilhussain5057@gmail.com

2nd Inam Ul Haq
Dept of Computer Science Engineering
Chandigarh University
Punjab, India
mirinam525@gmail.com

3rd Malik Muzamil Ishaq
Dept of Computer Science Engineering
Chandigarh University
Punjab, India
muzamilmalik705@gmail.com

4th Navjot Singh Talwandi
Dept of Computer Science Engineering
Chandigarh University
Punjab, India
navjotsingh49900@gmail.com

5th Vikash Yadav
Dept of Computer Science Engineering
Government Polytechnic Bighapur Unnao
Uttar Pradesh, India
vikas.yadav.cs@gmail.com

6thKaran Sarawagi
Dept of Computer Science Engineering
Chandigarh University
Punjab, India
kanuagarwal01@gmail.com

***Abstract*—Deepfake technology has been undoubtedly growing at a rapid pace since 2017. Specifically, GAN model and other techniques related to it were majorly popularized at that time. The recent trends of emerging artificial intelligence and deep learning models make it easier for the forgers to create fake images, audio, and videos. Although manipulation of visual and auditory media is very traditional method of fooling someone, the entrance of deepfakes has marked a turning point in the creation of fake content. Further advances in AI and machine learning, they offer automated procedures to create fake content that is much harder for common people to differentiate from the original content. As Deepfake technology advances, so do the techniques to detect them like GAN's (Generative Adversarial Networks). Other deep learning approaches are being developed to counteract GAN-generated deep fakes. This paper particularly looks at the research done in this area by various other researchers and engineers. It also focusses on research that can influence it by looking at papers regarding human pose transfer, human motion transfer, and human motion generation. Finally, the paper highlights the promising directions and future research opportunities in the field of deepfake detection. As the arms race between deepfake generators and detectors continues, collaborative efforts from academia, industry, and policymakers are crucial to developing robust defenses against the misuse of deepfake technology**

***Index Terms*—Deepfake, AI, ML, GAN's, Deep learning, human motion transfer.**

I. Introduction

In the era of advanced artificial intelligence and deep learning, the term "deepfake" has emerged as both a technological marvel and a growing concern. In the past two decades, the trend of creating fake media using deepfake has increased rapidly. This makes use of deep learning models, particularly GAN's. "The 21st centuries' answer to photoshopping, deepfakes use a form of artificial intelligence called deep learning to make images of fake events, hence the name deepfake (deep learning + fake)." Some of the wrong uses of deepfakes are pornography (with the image of popular actresses), political speeches where the audio of a particular person can be used over the face/video of another person using the lip-syncing technology or similar thing can be done for religious speeches. Deepfakes were first introduced into the world when some adult video clips were tampered with the face of a popular personality. After this incident, the reddit community had an explosion of fake content. Another instance is when former president Barack Obama's speech was tampered with. The primary concern surrounding deepfakes lies in their potential for fraud. This deception can manifest in various ways, from spreading false information and hoaxes to perpetrating identity theft and financial fraud. As a result, deepfakes have garnered attention not only for their creative potential but also for their capacity to undermine trust and manipulate public discourse.

In an age where the boundaries between reality and fiction are increasingly blurred, understanding deepfakes is essential for society to navigate the challenges and opportunities they present.

II. Literature Review

N Choi et al (2023) with no requirement for artificial intelligence (AI), this research provides a ground- breaking blockchain-based technique for identifying online movie fraud. It responds to the expanding demand for impartial and reliable detection methods in the face of developing video content by combining user contributions and cutting-edge algorithms to enable transparent and reliable detection of fake movies. In comparison to public blockchains built on PoW, Hyperledger Fabric offers efficiency benefits for service provider agreements. Significant study was done on smart contract design and testing by Sánchez-Gómez et al. and Górski, who emphasized the value of methodical consideration and verification techniques throughout the software development life cycle. [1]

Rajneesh Rani et al (2022) discussed that deepfake is a method for producing fake faces that may be used in place of real ones in photos or movies. It is based on machine learning and artificial intelligence. General Adversarial Networks (GANs) are the key element that makes deep fakes more realistic than ever; with the aid of these network systems incredibly convincing deep fakes can be formed. Although many different strategies have been developed in this field, the majority of them fall under the headings of facial artifacts, neural networks, and 3D head position. [2]

B Xu et al (2021) discussed that Deep learning is currently used to identify Deepfake films, which build deep neural networks to identify the frame sequence following framing. Recurrent Neural Network (RNN)-based detection is suggested in [9] based on the observation that some inconsistent choices, such as illuminants, exist across scenes containing fake frames. Convolutional Neural Network (CNN) is used to extract face information after first dividing the movie into frames. Long Short-Term Memory (LSTM) network is a collection of features from the videos' eye blinking detection and sent into an LSTM network to identify the subjects' eyes. The results of the experiments demonstrate that the suggested method may successfully identify Deepfake films.[3]

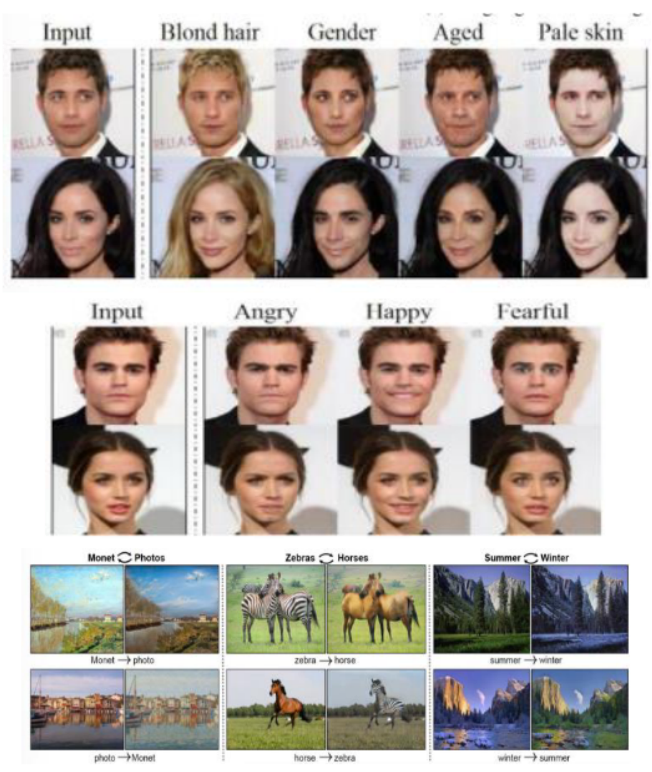

Fig. 1. Examples of images generated using GAN

Tao Zhang et al (2022) demonstrated that recent improvements in deepfake production make it more realistic, which makes it challenging to detect. The term "facial reenactment" refer to realistic-looking but false images, sounds, and movies produced by AI techniques. Deep fake has posed a serious danger to national security, democracy, society, and privacy, necessitating the use of existing datasets and detection techniques. Deepfake identification presents some difficulties, such as a dearth of good standards and datasets. [4]

Lakshmanan Nataraj et al (2019) discussed that deepfakes, image-to-image translations, and other automated techniques using GAN have grown incredibly popular. In order to detect GAN generated false images, a mix of co-occurrence matrices and deep learning has been used in this paper. In the pixel domain, co-occurrence matrices are collected for the three color channels, and a deep convolutional neural network (CNN) architecture is used to train a model. The use of GANs in steganalysis to detect the existence of concealed data in an image has received considerable research. The co-occurrence technique has historically relied on a matrix and a machine-leaning classification algorithm like SVM. [5]

Joseph Bamidele Awotunde et al (2022) investigates privacy issues and the propagation of false information caused by the pervasive usage of deepfake technology, which creates fake videos. This research introduces a deepfake fake model that runs effectively by utilizing a five-layered Convolution Neural Network (CNN) technique. A novel approach to deep fake detection using a specialized CNN architecture strengthened by ReLU activation functions. In order to address new issues, it also acknowledges the dynamic nature of deepfake technology and emphasizes the urgent need for continual improvements in deepfake detection. [6]

Trisha Mittal et al (2020) discussed that main method used in this study to spot deep fakes is to take advantage of how the audio and visual components of a video source interact with one another. Prior studies, including those in psychology and multimodal machine learning, have repeatedly shown a high correlation between different modalities related to the same topic. There is a positive link between audio and visual modalities, a relationship that has been useful in the field of multimodal perceived emotion identification. This results from the various ways that emotions are expressed, which may cause our method to find discrepancies in the modalities of real videos and, as a result, incorrectly classify them by labeling them as false. [7]

Yuezun Li et al (2018) discussed that deepFake technology mainly relies on training deep neural networks with facial photos. In this research, we provide a revolutionary deep learning-based system that can distinguish between DeepFake films and real ones with great accuracy. Our method takes advantage of a distinctive feature of DeepFake videos: the DeepFake algorithm can only produce facial images of a limited size. As a result, these synthesized faces must go through an affine warping procedure to match the source face's arrangement. Because the warped face area and the surrounding context have different resolutions, this warping procedure generates obvious artifacts. These artifacts work as helpful markers for spotting DeepFake films. By employing a customized Convolutional Neural Network (CNN) model to compare the generated facial regions with their surrounding areas, our method can detect these artifacts.[8]

Y Li et al (2019) discussed that we provide Celeb- DF3, a brand-new, sizable dataset created particularly to test and analyse DeepFake detection algorithms. More than 2 million frames from 5,639 DeepFake videos make up this dataset. These videos are contrasted with true source videos that came from publicly accessible sources. 59 celebrity YouTube videos reflecting a range of ages, genders, and ethnicities. When compared to previous datasets, the visual quality of the DeepFake movies in Celeb-DF3 is noticeably better. The findings highlight the substantial obstacle Celeb-DF3 presents to most current detection techniques. [9]

T Liang et al (2020) discussed a novel method for efficiently capturing a variety of face emotions, head positions, and complex situations in movies using discrete, large interval sampling. This approach seeks to reach thorough conclusions about a video's veracity, especially in situations when real frames are intermingled with edited content. We introduce the SDHF, a hierarchical framework that uses 2D convolutional neural networks to extract frame-level data. Then, clip-level and video-level features are extracted using a 1D convolutional aggregator. SDHF Framework combines elements at three separate levels—frame, clip, and video—allowing for a thorough evaluation of video authenticity. The Best Sampling Method enables the thorough sampling of video situations. Assessment Using Multiple Datasets evaluates the SDHF approach's effectiveness using the DFDC, Celeb-DF, Face-Forensics++, and UADFV datasets.[10]

III. Methodology

A. Entire face synthesis

It aims to generate non-existent fake face image xf from random vector v with neural network (·). That is xf = (v). GANs and VAEs are both relevant neural networks for full-face synthesis tasks. Some of the popular entire face synthesis techniques are PGGAN, StyleGan, etc. These techniques have been proven to produce high-quality deepfake images. The VAEs produce rather blurred images 4 as compared to the GANs. The main reason behind these blur images is that the training principle makes VAEs assign a high probability. As the images produced by VAEs are not realistic enough, this method specifically introduces GAN related works. GANs are used to learn the mapping from a random distribution (usually Gaussian noise) to a distribution of human faces. They aim to generate realistic images that are difficult to distinguish from real human faces

B. Attribute manipulation definition

It aims to modify the facial properties P of a real face image xr to generate a new fake image xf with neural networks (·, ·). That is xf = (xr, P). With a better attribute disentangle technique, the GANs for attribute manipulation

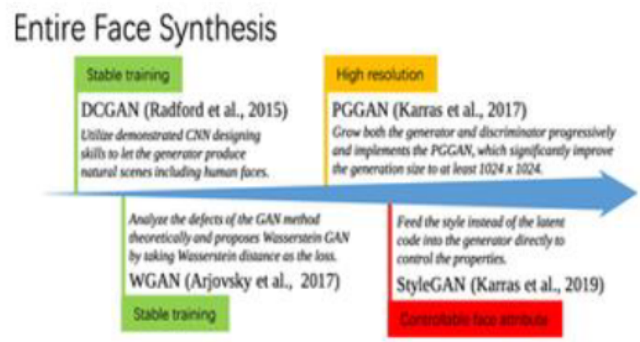

Fig. 2. Entire face synthesis[14]

can achieve more accurate attribute control. Existing state-of-the art methods [e.g., HifaFace (Gao et al. 2021d)] can perform accurate face editing while maintaining rich details of non-editing areas.The classical examples are StarGAN (Choi et al. 2018) and selective transfer GAN (STGAN) (He et al. 2019b). Invertible conditional GAN (IcGAN) (Perarnau et al. 2016) is the earliest attempt in GAN- based facial attribute manipulation.

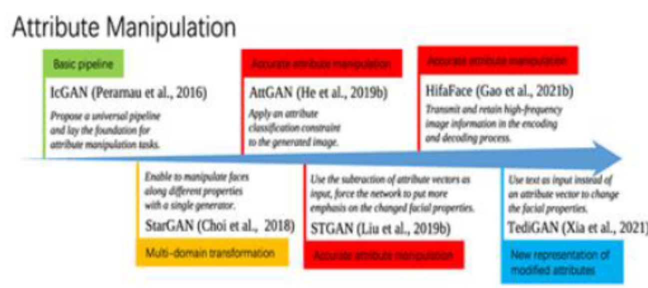

Fig. 3. Attribute Manipulation[14]

C. Identity swap definition

Identity swap aims to replace the identity of source image xs by the identity ti of target image xt with neural network (·, ·) and generate a new fake image xf. That is xf = (xs,ti)

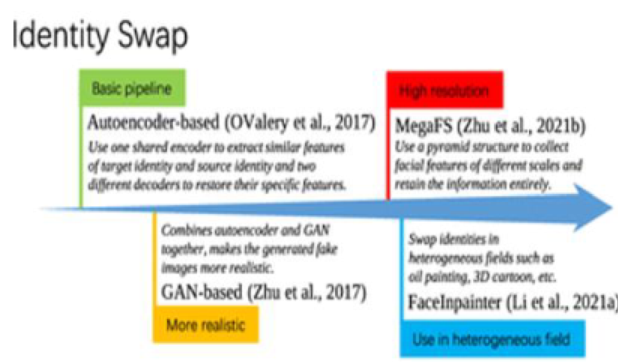

Fig. 4. Identity swap[14]

IV. PROPOSED SOLUTION RESULT AND ANALYSIS

As discussed earlier, GAN's are widely used technologies in making deepfakes and they are one of the most accurate models available till now. Images and videos created by GAN's are closely related to the original media. Since there is no 100accurate model to detect deepfakes yet, a combination of two models is generally preferred. In this paper, we will be working on DCGAN (an architecture of GAN) and CNN.

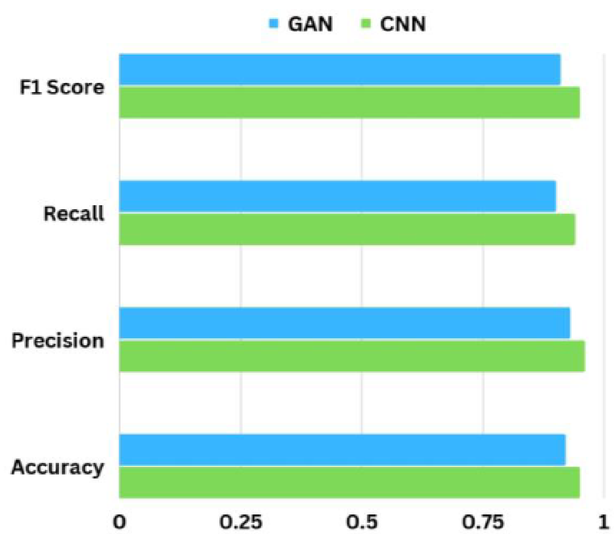

Fig. 5. Identity swap[14]

Step 1: The initial stages include the pre-processing of the data. This step involves importing a dataset of images and videos into a framework. Then the data is divided into two parts: training and testing set to train and assess the model respectively. Moreover, resizing the images to a standard size to improve the accuracy of the model is preferred.

Step 2: Next step is to create two models: the generator to generate fake images and the discriminator to evaluate real and fake images. Now, the DCGAN model is trained by feeding it with real and convincing fake images. DCGAN loss function: It measures how well the discriminator can discern between authentic and fraudulent images and how well the generator can produce realistic images. Reduce this loss as much as possible to raise the caliber of photographs that are generated. DCGANloss= Gloss+Dloss, (G is generator and D is discriminator)

Where, Gloss = -log(D(G(z)))

Dloss = -log(D(x))-log(1-D(G(z)))

(x is real image and z is noise vector)

Step 3: Now that the generator model can produce fake images and it is saved now it's time to define CNN model to extract features from images. The CNN model is trained using both the fake images created by DCGAN generator and real images from the loaded dataset. It trains the model to distinguish between them. After this, CNN model extracts features from images used for image classification. CNN loss function measures how well CNN model performs in distinguishing.

CNNloss = -Y*log(p)-1(1-Y)*log(1-p)

X is input image, Y is generated image and p is predicted probability

Step 4: In the final step of preparation of the model, we now combine the features extracted by CNN and DCGAN model. The goal of this information fusion is to make use of both the CNN's feature learning skills and the DCGAN's image generating capabilities. Now, the combined features are served as an input to a new layer of classifier. This layer makes classification prediction. The model is trained with these combined features to recognize patterns. CNN combined loss function quantifies the performance of the combined model in its classification task.

V. Conclusion

The conclusion of this research paper summarizes what is deepfake - "The 21st centuries' answer to photoshopping, deepfakes use a form of artificial intelligence called deep learning to make images of fake events, hence the name deepfake (deep learning + fake)." the various ways in which deepfakes are used to create and spread forged media on internet and how it is exploited in different ways like pornography, religious and political speeches. We have also discussed different research and study that has been conducted on deepfakes and methods to identify them. In this paper, we have also created the model using CNN and GAN to identify deepfakes and discussed about its results and analysis of features and their limitations.

VI. FUTURE WORK

Although, Deepfakes have been into the market since almost two decades, but still there is a lot of research work to be done on this topic as we do not have any accurate foolproof method to detect manipulated media using deepfakes. Some CNN and machine learning algorithms have proved to be very helpful in achieving our target but none of them is 100reliable. So, here are some of the works that can be done:

1. Development of more advanced and accurate deepfake detection method. This might involve the use of more sophisticated machine learning algorithms, combining multiple modalities (e.g., audio and video), and exploring novel features or representations.

2. Researchers can focus on making these models more transparent and understandable, helping users trust the decisions made by these systems.

3. Deepfake technology is extending beyond images and videos to generate text-based content. Future work can delve into detecting fake text generated by GANs to combat disinformation and fake news.

4. Developing real-time deepfake detection systems is essential, especially in applications where immediate action is required, such as in live-streamed content or video conferencing.

5. Deepfake techniques are continually evolving. Researchers should work on developing detection methods that can generalize to detect deepfakes created using new technologies and approaches.

6. Creating diverse and comprehensive datasets for deepfake detection is crucial. Future research can focus on curating more challenging and realistic datasets for benchmarking detection methods.

7. Developing user-friendly tools that enable individuals to check the authenticity of media content easily can empower users to protect themselves from deepfake-related threats. Future work in the field of deepfakes should not only focus on detection but also encompass prevention, mitigation, and the ethical considerations surrounding this technology. It requires a multidisciplinary approach that combines expertise in machine learning, computer vision, cybersecurity, law, ethics, and public awareness

References

[1] Choi, N. and Kim, H., 2023. DDS: Deepfake Detection System through Collective Intelligence and Deep-Learning Model in Blockchain Environment. Applied Sciences, 13(4), p.2122.

[2] Rani, R., Kumar, T., Sah, M.P. (2022). A Review on Deepfake Media Detection. In: Sharma, H., Shrivastava, V., Kumari Bharti, K., Wang, L. (eds) Communication and Intelligent Systems . Lecture Notes in Networks and Systems, vol 461. Springer, Singapore.

[3] Xu, B., Liu, J., Liang, J., Lu, W. and Zhang, Y., 2021. DeepFake Videos Detection Based on Texture Features. Computers, Materials Continua, 68(1).

[4] Zhang, T. Deepfake generation and detection, a survey. Multimed Tools Appl 81, 6259–6276 (2022). https://doi.org/10.1007/s11042-021-11733-y

[5] Nataraj L, et al (2019) Detecting GAN generated fake images using co-occurrence matrices. arXiv:1903.06836

[6] Awotunde, J.B., Jimoh, R.G., Imoize, A.L., Abdulrazaq, A.T., Li, C.T. and Lee, C.C., 2022. An Enhanced Deep Learning-Based DeepFake Video Detection and Classification System. Electronics, 12(1), p.87.

[7] Mittal, T., Bhattacharya, U., Chandra, R., Bera, A. and Manocha, D., 2020, October. Emotions don't lie: An audio-visual deepfake detection method using affective cues. In Proceedings of the 28th ACM international conference on multimedia (pp. 2823- 2832).

[8] Li, Y. and Lyu, S., 2018. Exposing deepfake videos by detecting face warping artifacts. arXiv preprint arXiv:1811.00656. 7

[9] Li, Y., Yang, X., Sun, P., Qi, H. and Lyu, S.C.D., 1909. A Large-scale Challenging Dataset for DeepFake Forensics (2019). URL http://arxiv.org/abs/1909.12962, 35, p.36.

[10] Liang, T., Chen, P., Zhou, G., Gao, H., Liu, J., Li, Z. and Dai, J., 2020, November. SDHF: spotting deepfakes with hierarchical features. In 2020 IEEE 32nd international conference on tools with artificial intelligence (ICTAI) (pp. 675-680). IEEE.

[11] Liy, C.M. and InIctuOculi, L.Y.U.S., 2018, December. Exposingaicreatedfakevideosbydetectingeyeblinkin g. In Proceedings of the 2018 IEEE International workshop on information forensics and security (WIFS), Hong Kong, China (pp. 11-13).

[12] Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A. and Schulman, J., 2022. Training language models to follow instructions with human feedback. Advances in Neural Information Processing Systems, 35, pp.27730-27744.

[13] Baltrušaitis, T., Ahuja, C. and Morency, L.P., 2018. Multimodal machine learning: A survey and taxonomy. IEEE transactions on pattern analysis and machine intelligence, 41(2), pp.423-443.

[14] Baccarella, C. V., Wagner, T. F., Kietzmann, J. H., McCarthy, I. P. (2018). Social media? It's serious! Understanding the dark side of social media. European Management Journal, 36(4), 431-438.

[15] Agrawal, A., Gans, J.S., Goldfarb, A., 2019. Exploring the impact of artificial intelligence: prediction versus judgment. Inf. Econ. Policy 47, 1–6. https://doi.org/10.1016/j. infoecopol.2019.05.001.

[16] Amabile, T.M., 2020. Creativity, artificial intelligence, and a world of surprises. Acad. Manag. Discov. 6 (3), 351–354. https://doi.org/10.5465/amd.2019.0075.

[17] Appio, F.P., Frattini, F., Petruzzelli, A.M., Neirotti, P., 2021. Digital transformation and innovation management: a synthesis of existing research and an agenda for future studies. J. Prod. Innov. Manag. 38 (1), 4–20. https://doi.org/10.1111/jpim.12562.

[18] Barrett, R., 2020. Innovation precinct transforming industry. The Australian, 12 November. Available at: https://www.theaustralian.com.au/special-reports/innova tion-precinct-transforming-industry/news-story/846eb69184f6900b3cdc7e2dd4e bc406.

[19] Benner, M.J., Tushman, M.L., 2015. Reflections on the 2013 Decade Award – "exploitation, exploration, and process management: the productivity dilemma revisited" ten years later. Acad. Manag. Rev. 40 (4), 497–514. https://doi.org/ 10.5465/amr.2015.0042.

[20] Bogers, M.L.A.M., Garud, R., Thomas, L.D.W., Tuertscher, P., Yoo, Y., 2021. Digital innovation: transforming research and practice. Innovation 1–9. https://doi.org/ 10.1080/14479338.2021.2005465.

[21] Brynjolfsson, E., McAfee, A., 2014. The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. WW Norton Company, New York, NY.

[22] Brynjolfsson, E., Mitchell, T., 2017. What can machine learning do? Workforce implications. Science 358 (6370), 1530–1534. https://doi.org/10.1126/science. aap8062.

[23] Brynjolfsson, E., Rock, D., Syverson, C., 2017. Artificial intelligence and the modern productivity paradox: a clash of expectations and statistics. National Bureau of Economic Research Working Paper Series 24001. https://doi.org/10.3386/w24001.

[24] Bughin, J., Hazan, E., Ramaswamy, S., Chui, M., Allas, T., Dahlstr om, P., Henke, N.,

[25] Trench, M., 2017. Artificial Intelligence: The next Digital Frontier? McKinsey Global Institute. Available at. https://www.mckinsey.com/ /media/mckinsey/industries/advancedanscussion-paper.ashx.

[26] Burstr om, T., Parida, V., Lahti, T., Wincent, J., 2021. AI-enabled business-model innovation and transformation in industrial ecosystems: a framework, model and outline for further research. J. Bus. Res. 127, 85–95. https://doi.org/10.1016/j. jbusres.2021.01.016.

[27] Ceipek, R., Hautz, J., De Massis, A., Matzler, K., Ardito, L., 2021. Digital transformation through exploratory and exploitative internet of things innovations: the impact of family management and technological diversification. J. Prod. Innov. Manag. 38 (1), 142–165. https://doi.org/10.1111/jpim.12551.

[28] Chalmers, D., MacKenzie, N.G., Carter, S., 2021. Artificial intelligence and entrepreneurship: implications for venture creation in the Fourth Industrial Revolution. Entrep. Theory Pract. 45 (5), 1028–1053