From Mind-Reading to Health-Reading Machines: Towards Contactless Health Diagnosis using Generative Artificial Intelligence

Abdenour Hadid
Professor of Chair of Excellent in Artificial Intelligence,
Sorbonne Center for Artificial Intelligence, Sorbonne University Abu Dhabi, Abu Dhabi,
UAE,
abdenour.hadid@sorbonne.ae

Abstract- Generative Artificial Intelligence (AI) has recently shown impressive results across a wide spectrum of tasks. Inspired by the emerging notion of mind-reading machines and by the medical evidences indicating that some diseases and brain disorders produce facial abnormalities and interrupt normal facial expression formation, we are 1) establishing, in close interaction with medical practitioners, the basic understanding of the correlations between visual patterns and health conditions, and 2) developing novel computational models, based on Generative AI, to learn the identified correlations. Devising computational models for detecting abnormalities reflective of internal diseases based only on visual information is a highly intriguing research problem. The goal is to achieve ground-breaking results in unobtrusive health diagnosis and to gain insight into the correlation between face, body, and mind. This work has the potential to open new research directions in several fields and to bridge the gap between the research activities conducted in several disciplines like engineering and medicine. This article explains the main idea behind contactless health reading using generative artificial intelligence.

Keywords- Generative Artificial Intelligence; Health Diagnosis; Computer Vision; Deep Learning.

1.   Introduction

Inspired by generative artificial intelligence together with medical evidences indicating correlation between facial symptoms and some internal diseases (e.g., [1-4]), our work aims to devise computational models for detecting abnormalities reflective of diseases in person's facial structures and expressions based only on visual information. This would help designing futuristic unobtrusive technologies for health diagnosis and monitoring that people can effortlessly use in their daily lives without any contact. Imagine a "magic" mirror at home which unobtrusively monitors your physiological health measurements (e.g., heart rate and blood pressure), recognizes your affect states (e.g., stress and fatigue) and diagnoses possible diseases (e.g., imminent stroke or kidney infection) by only observing your face while you are in front of the mirror for activities such as shaving, brushing your teeth, or washing your face. Such a "magic" medical mirror could, for instance, provide real-time feedback information about your health condition and even act as an assistive and therapeutic device by displaying a realistic 3D face avatar for engaging emotional interaction and preventive activities. As a second example, imagine a doctor wearing "smart" glasses which can unobtrusively estimate the

pain of neonates and post-surgery patients that are incapable of articulating and expressing their pain experiences. During the past decade, there have been numerous research and development efforts in the field of wearable health-monitoring systems that were motivated by the need to monitor a person's health status outside of the hospital [5-6]. However, most current techniques for health monitoring typically require users to strap on bulky sensors, chest straps or sticky electrodes. This obviously discourages regular use because the sensors can be uncomfortable or encumbering. Hence, to make health monitoring part of the fabric of everyday life, we believe that there is a need for new technologies that are comfortable (e.g., non-invasive and contact free), simple to use and unobtrusive. The idea of using computer vision for non-contact healthcare and wellness has the promise to improve the future of healthcare by enabling proactive and unobtrusive personal health, mental and activity management. It would enable long-term daily health monitoring at home, which is important for the treatment and management of many chronic illnesses, neurological disorders, and mental health issues such as diabetes, hypertension, asthma, autism spectrum disorder, depression, drug addiction, etc.

On the other hand, the development of generative AI models is already raising some concerns about privacy preservation and the potential emergence of super intelligent systems without adequate safeguards. Generative AI can indeed reveal sensitive health data or unintentionally produce biased or incorrect information due to biases in its training data. Moreover, generative models usually require large amount of high-quality unbiased data to operate - which is lacking in medical domain. Other equally important issues include the latency for generating high-quality samples and the massive computing power which is needed to train generative models. All these issues are important to be taken into consideration to effectively capitalize and guarantee beneficial use of generative AI.

The aim of this article is to explain the main potential and challenges behind contactless visual health reading using generative artificial intelligence, and to point out some future directions.

2. Motivations

Faces can be seen as projector panels of the mechanisms that govern our emotional behaviors and health (see Figure 1 for potential health conditions that can be reflected from a face). Automatic analysis of face patterns and movements (e.g., in terms of the activity of the facial muscles causing the visible changes in facial expression) for medical monitoring and diagnosis is indeed an exciting area of research. For instance, researchers at Massachusetts Institute of Technology (MIT) have published the pioneer work [6] pointing out the possibility of measuring heart rate, respiratory rate, and heart rate variability from face images acquired by a webcam under controlled settings. Moreover, researchers from the University of Southern California have shown, in preliminary investigations published in the prestigious Journal of Neurology, that eye movements could help in diagnosis of neurological disorders such as Parkinson's disease [1].
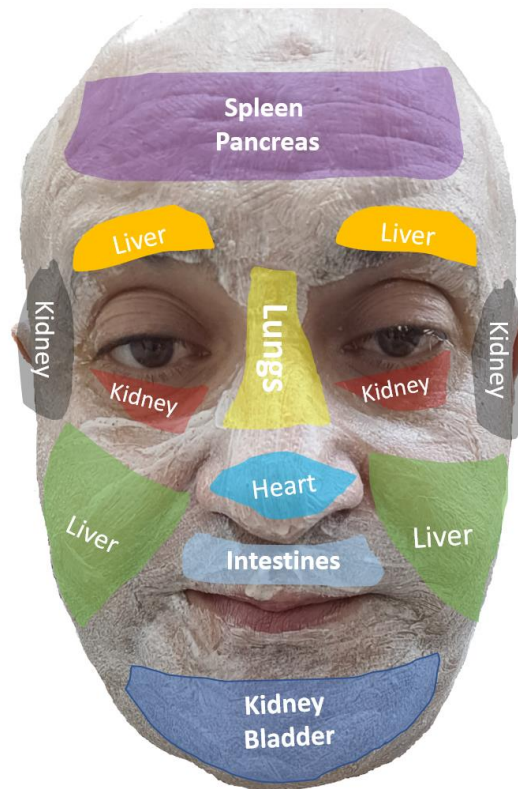
Figure 1. Several internal diseases may have face symptoms.

Moreover, children with autism usually share some similar facial features (such as wider eyes, a wider mouth area and a broad upper face area but a shorter middle region of the face with a flattened nasal bridge). These signs are so subtle that they may not be noticeable for a human eye. They are however computationally very distinguishable. So, tools for early diagnosis of children with autism are important as they can improve a child's chances for a successful therapy. Many other diseases also produce facial abnormalities and interrupt normal facial expression formation (Figure 1). For instance, facial asymmetry is one of a number of traits associated with stroke problems; a white ring bordering the colored part of the eye can be a sign of fatty deposits caused by high cholesterol; eye twitching, also known as myokymia, is usually caused by random muscle contractions and may be triggered by lack of magnesium; a drooping eyelid can be a sign of a stroke, lung cancer, eye strain or stress; cracked lips are sometimes caused by anemia due to iron deficiency and may also be a sign of diabetes and so on. Furthermore, chronic affect states such as pain, stress, anxiety, and depression can also carry a wide array of pathophysiological risks, including cardiovascular disease, cerebrovascular disease, diabetes, and immune deficiencies, to cite only a few. New unobtrusive technologies that automatically recognize such states from the facial patterns and movements can be extremely powerful, both diagnostically and therapeutically. As a diagnostic tool, the technology could help individuals and clinicians gaining insight into the conditions that consistently provoke such states. As a therapeutic tool, the technology could be used to automatically initiate interventions. Inference of user's complex mental states (including emotions, cognitive states, intentions, beliefs, desires and focus of attention) from nonverbal cues and observed behavior is currently the challenge of the ongoing efforts toward mind-reading machines.

We aim to lay the core foundations for using computer vision and generative AI in healthcare by 1) establishing, in close interaction with medical practitioners, the basic understanding of the correlations between facial visual symptoms and health conditions, and 2) developing novel computational models that encode the identified correlations. While our primary focus is on health measurements that can be reflected from the face patterns and movements, other body parts and behavioral characteristics such as gait and gestures can also be considered in longer-term investigations, as such cues also convey health signals.

Our efforts are focusing on developing novel AI models for futuristic technologies that will enable us to continuously monitor our health and track our progress in all stages of life and diseases. The idea of using computer vision along with generative AI is appealing as it has the promise of enabling proactive and unobtrusive personal health management.

3.   Challenges

Detecting abnormalities reflective of internal diseases in patient's facial structures and expressions based only on visual information is a highly intriguing and under-explored research problem. The critical issues concern the establishment of basic understanding of the correlations between facial visual symptoms and health conditions and then the development of new AI models that encode the identified correlations. We need to devise algorithms that would enable the construction of AI models from data representing health measurements and physiological processes occurring at different time scales and across a wide range of people and ages. The fundamental challenge of vision-based face analysis for healthcare partially lies in the enormous richness of the face patterns and the high complexity of the facial movements, e.g., due to no-rigidness. This raises several fundamental questions such as: how computer vision systems can differentiate between signs of diseases and make-ups and disguises, based only on visual information? which kinds of features (e.g., color, texture, motion, or depth) are most effective to diagnose a given disease? what facial regions (e.g., lips or eyes) are associated to what diseases? which kinds of data sensing technologies (e.g., thermal or 3D imaging) are necessary for detecting a given anomaly? how to efficiently infer complex mental states (e.g., stress) from subtle facial changes? can machine vision transcend human abilities in visual health diagnosis? how to enable computer vision to interact best with medical science? how to identify enough number of patients with particular disease for constructing representative and balanced datasets? Can generative AI help in creating synthetic samples for training AI models? how generative AI can be used while preserving the privacy? what is the degree and nature of the correlation between face, body, and mind? All of these are challenges that require thorough exploration.

Moreover, while generative AI has shown immense potential in various domains, its resource-intensive nature may hinder real-time usage and scalability. In fact, the large-scale generative AI models require significant computational resources and electrical power to operate, resulting in elevated energy consumption and significant carbon emissions. Consequently, this may restrict their deployment in real-world applications. Future efforts are then needed to design novel and efficient architectures capable of generating high-quality samples in real-time, vital for constrained platforms like in edge computing.

As generative artificial intelligence models can be considered in their early stages, here are some of the key challenges that still must be tackled to ensure applicability in healthcare:

3.1. Handling Sensitive Data

Health data refers to information that relates to the health status of a person. This is considered as highly sensitive data and is subject to additional precautions and strict rules [7]. AI models should incorporate adequate measures beyond pseudonymization and encryption.

3.2. Data Scarcity

Among the main challenges in using AI in healthcare is the lack of enough samples for training the models. Generative AI models usually require a large amount of high-quality and unbiased data to operate. Although generative AI models can be used to produce synthetic data for training, other strategies can be devised for scenarios with limited data [8].

3.3. Imbalanced Classes

A prevalent issue encountered in healthcare applications is the "class imbalance" problem, signifying disparities in the frequencies of positive samples (samples with a given disease) compared to the negative samples (samples without the given disease). Developing data-driven AI models from imbalanced datasets require additional attention [9].

3.4. High Computational Power

Large-scale computing infrastructures are usually necessary to maintain and develop generative models [10]. For example, diffusion models could require millions or billions of images to train. To train such large datasets, massive computing power (clusters with hundreds of GPUs) maybe be needed. Moreover, due to the very large-scale of generative AI models, there may be latency present in the time it takes to generate a sample.

4.   Directions

4.1. Multimodal Analysis

Visual symptoms, physiological signals, affect states and some internal diseases are correlated and should be studied in a multimodal framework [11]. For instance, a high blood pressure under stress emotional state should not be considered as a serious warning. Of importance is also the combination of complementary information provided by different sensor data.

4.2. Collaboration Between Biomedical and Computer Sciences

Although generative AI can offer elegant solutions to many challenges on healthcare, it is essential to acknowledge that AI is not a complete solution that can replace human experience. The aim of AI is to support the clinicians as many problems necessity a through comprehension of the underlying medicine. This calls for a multi-disciplinary approach at the frontiers of computer vision, medicine, neuroscience, and psychology. The close collaboration between clinicians and researchers in all these disciplines is essential.

4.3. Sensing Beyond the Visible Light

Visual images acquired using conventional 2D cameras and/or at visible spectrum may have inherent restrictions that hinder the inference of some specific health details in the visual world [12]. One promising approach for dealing with those limitations is using images acquired beyond the visible spectrum and/or using non-conventional imaging (e.g., depth).

4.4. Data Efficient Modeling

Among the main challenges in using AI in healthcare is the lack of enough samples for training the models. Generative AI models usually require a large amount of high-quality and labeled data to operate. Although generative AI models can be used to produce synthetic data for training, other strategies can also be devised for scenarios with limited data. This includes few-shot learning [13], transfer learning [14], and domain adaptation [15] which offer the potential to enhance the AI performance when data is scarce.

4.5. Federated Learning and Privacy-Preserving Modeling

To alleviate concerns about patient privacy, federated learning is often suggested [16]. Federated learning is a collaborative approach to train machine learning models, not requiring an exchange of data from client devices (e.g., hospital) to global servers. The raw data on edge devices can be used to train a model locally, increasing data privacy. Then, each edge device shares its model's parameters to "reconstruct" the final model in the server.

4.6. Explainable and Trustworthy Modelling

Lack of transparency is among the barriers to ubiquitous use of AI in healthcare [17]. To strengthen trust, people need to understand how the AI technology works. Clinicians should indeed be confident in the AI systems. A possible step towards trustworthy AI is to develop explainable AI. Explainable AI aims to design models capable of generating decisions that a human could understand and interpret. This will strengthen the trust in AI models.

5.   Conclusion

The integration of generative AI promises a paradigm shift in healthcare. However, with every innovation comes skepticism. The development of generative models in healthcare may indeed raise concerns about privacy protection and the potential emergence of super-intelligent machines without adequate safeguards. Generative AI can unintentionally reveal sensitive data or produce incorrect information due to biases in the training data.

Moreover, generative models usually require large amount of high-quality, unbiased data to operate. Other issues concern the latency for generating high-quality samples and the massive computing power which is needed to train generative models. So, at the end, can we really trust the outcomes of generative AI models for critical operational decision in healthcare? This call for explainable and trustworthy generative AI research.

In conclusion, generative AI will surely play a pivotal role in the future of healthcare. It is a powerful tool that can transform the way we interpret and understand healthcare data. While there are challenges to overcome, the potential benefits far outweigh the hurdles.

Acknowledgement

References

[1] Tseng, P., Cameron, I. G. M., Pari, G., Reynolds, J. N., Munoz, D. P., and Itti, L. High-throughput classification of clinical populations from natural viewing eye movements. *Journal of Neurology 260 (Jan 2013)*, 275-284.

[2] J. Thevenot, M. B. López and A. Hadid, "A Survey on Computer Vision for Assistive Medical Diagnosis from Faces," in *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 5, pp. 1497-1511, Sept. 2018, doi: 10.1109/JBHI.2017.2754861.

[3] N. Kour, Sunanda and S. Arora, "Computer-Vision Based Diagnosis of Parkinson's Disease via Gait: A Survey," in *IEEE Access*, vol. 7, pp. 156620-156645, 2019, doi: 10.1109/ACCESS.2019.2949744.

[4] M. Rajnoha, J. Mekyska, R. Burget, I. Eliasova, M. Kostalova and I. Rektorova, "Towards Identification of Hypomimia in Parkinson's Disease Based on Face Recognition Methods," *2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Moscow, Russia, 2018, pp. 1-4, doi: 10.1109/ICUMT.2018.8631249.

[5] Pantelopoulos, A., and Bourbakis, N. G. "A survey on wearable sensor-based systems for health monitoring and prognosis". *IEEE Transactions on Systems, Man, and Cybernetics*, Part C 40, 1 (Jan. 2010), 1-12.

[6] Poh, M.-Z., McDuff, D., and Picard, R. W. "Advancements in noncontact, multiparameter physiological measurements using a webcam". *IEEE Transactions on Biomedical Engineering* 58, 1 (2011), 7-11.

[7] Murdoch, B., "Privacy and artificial intelligence: challenges for protecting health information in a new era". *BMC Med Ethics* 22, 122 (2021). https://doi.org/10.1186/s12910-021-00687-3

[8] Alzubaidi, L., Bai, J., Al-Sabaawi, A. *et al.* "A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications". *J Big Data* 10, 46 (2023). https://doi.org/10.1186/s40537-023-00727-2

[9] Johnson, J.M., Khoshgoftaar, T.M. "Survey on deep learning with class imbalance". *J Big Data* 6, 27 (2019). https://doi.org/10.1186/s40537-019-0192-5

[10] Albert Reuther, Peter Michaleas, Michael Jones, Vijay Gadepally, Siddharth Samsi, Jeremy Kepner, "AI and ML Accelerator Survey and Trends" *2022 IEEE High Performance Extreme Computing (HPEC) Conference*, https://arxiv.org/abs/2210.04055

[11] Bayoudh, K., Knani, R., Hamdaoui, F. *et al.* "A survey on deep multimodal learning for computer vision: advances, trends, applications, and datasets". *Visual Computing* 38, 2939–2970 (2022). https://doi.org/10.1007/s00371-021-02166-7

[12] Berg, Amanda. (2019). "Learning to Analyze what is Beyond the Visible Spectrum". 10.3384/diss.diva-161077.

[13] Yaqing Wang, Quanming Yao, James Kwok, Lionel M. Ni, "Generalizing from a Few Examples: A Survey on Few-Shot Learning" https://arxiv.org/abs/1904.05046

[14] Abu Sufian, Anirudha Ghosh, Ali Safaa Sadiq, Florentin Smarandache, "A Survey on Deep Transfer Learning to Edge Computing for Mitigating the COVID-19 Pandemic", *Journal of Systems Architecture*, Volume 108, 2020.

[15] Garrett Wilson, Diane J. Cook, "A Survey of Unsupervised Deep Domain Adaptation", *ACM Transactions on Intelligent Systems and Technology,* Volume 11Issue 5Article No.: 51, pp 1–46, 2020

[16] Rieke, Nicola; Hancox, Jonny; Li, Wenqi; Milletarì, Fausto; Roth, Holger R.; Albarqouni, Shadi; Bakas, Spyridon; Galtier, Mathieu N.; Landman, Bennett A.; Maier-Hein, Klaus; Ourselin, Sébastien; Sheller, Micah; Summers, Ronald M.; Trask, Andrew; Xu, Daguang; Baust, Maximilian; Cardoso, M. Jorge (14 September 2020). "The future of digital health with federated learning". *npj Digital Medicine*. 3 (1): 119. https://arxiv.org/abs/2003.08119

[17] Aniek F. Markus, Jan A. Kors, Peter R. Rijnbeek, "The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies", *Journal of Biomedical Informatics*, Volume 113, 2021.