BMJ Open

Deep Learning-Based Facial Image Analysis in Medical Research: A Systematic Review Protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-047549
Article Type:	Protocol
Date Submitted by the Author:	01-Dec-2020
Complete List of Authors:	Su, Zhaohui; UT Health San Antonio, Center on Smart and Connected Health Technologies, Mays Cancer Center, School of Nursing Liang, Bin; Chinese Academy of Medical Sciences and Peking Union Medical College Shi, Feng; Shanghai United Imaging Intelligence Gelfond, J; University of Texas Health Science Center at San Antonio, Epidemiology and Biostatistics Wang, Jing; UT Health San Antonio
Keywords:	PUBLIC HEALTH, Information technology < BIOTECHNOLOGY & BIOINFORMATICS, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS, Telemedicine < BIOTECHNOLOGY & BIOINFORMATICS, BIOTECHNOLOGY & BIOINFORMATICS

Deep Learning-Based Facial Image Analysis in Medical Research: A Systematic Review Protocol

Zhaohui Su,[1*] Bin Liang,[2] Feng Shi,[3] Jonathan Gelfond,[4] Jing Wang[5]

1. Ph.D., M.A., School of Nursing, Center on Smart and Connected Health Technologies, Mays Cancer Center, UT Health San Antonio, San Antonio, Texas, 78229, U.S.A.; szh@utexas.edu

2. Ph.D., M.D., Department of Radiation Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 10010, China; leangbin@gmail.com

3. Ph.D., Department of Research and Development, Shanghai United Imaging Intelligence Co., Ltd., Shanghai 200232, China; feng.shi@united-imaging.com

4. M.D., Ph.D., Associate Professor, Department of Population Health Sciences, UT Health San Antonio, San Antonio, Texas, 78229, U.S.A.; gelfondjal@uthscsa.edu

5. Ph.D., M.P.H., R.N., FAAN, Professor, Vice Dean of Research, School of Nursing; Director, Center on Smart and Connected Health Technologies, UT Health San Antonio, San Antonio, Texas, 78229, U.S.A.; wangj1@uthscsa.edu

*Corresponding author. 7703 Floyd Curl Drive, San Antonio, TX, 78229, U.S.A.; szh@utexas.edu; ORCID ID: https://orcid.org/0000-0003-2005-9504

Deep Learning-Based Facial Image Analysis in Medical Research: A Systematic Review Protocol

Abstract

Introduction: Deep learning techniques are gaining momentum in medical research. Evidence shows that deep learning has advantages over humans in image identification and classification, such as facial image analysis in detecting people's medical conditions. While positive findings are available, little is known about patterns of utility of deep learning-based facial image analysis in the medical context. To address this gap, we aim to conduct a systematic review to identify the characteristics and effects of deep learning-based facial image analysis in medical research.

Methods and analysis: Databases including PubMed, PsycINFO, CINAHL, IEEEXplore, and Scopus will be searched for relevant studies published in English, using a search strategy that was developed in consultation with an academic librarian. Titles, abstracts, and full-text articles will be screened to identify eligible articles. A manual search of the reference lists of the included articles will also be conducted. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework was adopted to guide the systematic review process.

Conclusions: Insights gained from this systematic review will provide timely understanding of the characteristics, challenges, as well as opportunities in deep learning-based facial image analysis applied in the contexts of disease detection, diagnosis, and prognosis. In addition to gaining a connected and comprehensive understanding of the current application of facial image analysis, results of this review study will also be able to shed light on whether, similar to facial image analysis conducted in non-medical contexts, systematic bias and accuracy are present in medical research as well. Biased and inaccurate image analysis systems will not only exert unwarranted, though avoidable, disparities on patients (e.g., gender inequality), it may deprive some patients of potential benefits of deep-learning-based interventions for their health and wellbeing. Therefore, for the consideration of patients' welfare and the development of the practice, a timely understanding of the challenges and opportunities faced by research on deep-learning-based facial image analysis is needed.

Ethics and dissemination: NA

Study Protocol Registration: PROSPERO CRD42020196473

Keywords: facial image analysis; artificial intelligence; deep learning; convolutional neural network; abnormal facial expressions; facial analysis

Strengths and limitations of this study

- Deep learning is a mechanism that allows computers to solve complex problems by neural network architecture.
- Deep learning techniques are gaining momentum in medical research.
- Evidence shows that deep learning has advantages over humans in image identification and classification, such as facial image analysis in detecting people's medical conditions.
- While positive findings are available, little is known about patterns of utility of deep learning-based facial image analysis in the medical context.
- This study aims to conduct a systematic review to identify the characteristics and effects of deep learning-based facial image analysis in medical research.

A Systematic Review of Deep Learning-Based Facial Image Analysis in Medical Research: Insights and Implications

Introduction

The application of artificial intelligence (AI) in healthcare is gaining momentum [1-3]. AI, especially machine learning and deep learning techniques, is invigorating medical research on various fronts, from developing and deploying consumer-facing telemedicine tools, such as ePAL, an AI-powered smartphone application for patients to manage cancer-related pain [4], to the application of AI techniques in disease diagnosis and prognosis, such as using AI to detect and grade cancer in prostate biopsy samples [5]. Machine learning is a subset of AI that could be understood as methods that allow computers to learn by mimicking how humans learn (i.e., process data and learn from it), without being explicitly programmed [6]. One emerging machine learning approach, deep learning, has found to be particularly helpful in addressing medical issues, given its accuracy in image identification and classification [3, 7]. Deep learning micks the human brain by utilizing artificial neurons and layers (e.g., neural networks) [8].

While as presented in Figure 1, deep learning is a machine learning algorithm that allows computers to tackle complex problems via capitalizing on neural networks, such as convolutional neural networks (CNNs), that are rich in neurons, layers, and interconnectivity [9]. Simply put, deep learning is a mechanism that allows computers to solve complex problems by neural network architecture. This ability to develop complex network structures gives deep learning a distinctive advantage: it can automatically transform raw data input into meaningful features that enable pattern identification [3]. Deep learning technique has revolutionary potential in practical and research fields [10]. In practice, as deep learning effectively identifies objects, traffic signs, and faces, its adaptations have been widely applied in designing robots and self-driving cars [11-14]. Deep learning has also been widely adopted in biomedical and clinical research, particularly in the field of medical imaging [15-18].

---Insert Figure 1 here---

Medical conditions are often diagnosed by means of tests, such as measures like biopsy and diagnostic imaging. As diagnostic imaging is noninvasive and can facilitate personalized medicine, it is a preferred test option for patients and healthcare practitioners [19, 20]. This, in turn, has contributed to the exponential growth of medical imaging data and the increasing need for boosting medical image processing power to formulate diagnosis swiftly [20, 21]. Compared to traditional computer aided diagnosis for analyzing medical imaging, such as hand-crafted radiomics for tumor detection, deep learning methods are superior in their ability to process large quantities of medical images accurately and cost-effectively, without exerting a heavy workload on radiologists [22-26]. Evidence shows that deep learning-based medical image analysis was able to increase accuracy rates in various disease contexts, such as the identification of spinal disorder [27] and lung cancer histology [28], classification of skin lesion [29] and chronic gastritis [30], and the prediction of tumor-related genes [31] and vascular diseases [32].

As disease manifestations often show in various places in the human body, such as Down Syndrome can change patients' facial features, researchers have been investigating whether analyzing appearance features can facilitate early disease detection [1, 27, 33-35]. One promising field is deep learning-based facial analysis [2, 36, 37]. Applying the deep learning technique to perform facial recognition and analysis tasks, researchers found that the technique yielded superior results in identifying and classifying faces of people with cancer from those without [2]. Similarly, examining facial phenotypes of people with genetic disorders, findings indicate that the technique was effective and was able to yield an optimal 91% top-10 accuracy [38].

Evidence further indicates that, for some tasks involving identifying and classifying facial images, deep learning techniques have often performed on par or better than human beings [3, 35, 36, 39-41]. Comparing clinical and deep learning evaluations of microdeletion syndrome facial phenotypes, researchers found that deep learning outperformed clinical evaluations in terms of sensitivity and specificity by 96% [40]. These findings combined suggest that deep learning-based facial analysis technology has great potential to address complex medical challenges prevalent in healthcare. However, while useful insights are available, little is known about the current development of deep learning-based facial analysis utilized across different healthcare sectors. To date, no systematic review has investigated the state-of-the-art application of deep learning-based facial analysis in addressing medical issues. Therefore, to bridge this gap, we aim to systematically review the literature and present the characteristics and effects of deep learning-based facial analysis techniques applied in medical research.

Methods and analysis

This systematic review was registered with the International Prospective Register of Systematic Reviews database or PROSPERO (CRD42020196473) *a priori* to improve research rigor [42, 43]. The Principles of the Preferred Reporting Items for Meta-Analysis protocol (PRISMA) was adopted to guide this systematic review [44]. Our search strategy incorporated medical subject heading (MeSH) and keyword terms for the concept of deep learning and facial analysis. The search strategy was developed in consultation with an academic librarian, and subsequently will be deployed to target databases, including PubMed, PsycINFO, CINAHL, IEEEXplore, and Scopus (Table 1). The search will be initiated in September, 2020. Studies will be limited to journal articles published in English. We will adopt two additional search mechanisms to locate eligible articles: (1) a manual search of the reference list of the included articles will be performed, and (2) a reverse search of papers that cited articles included in the final review via Google Scholar. An academic librarian will facilitate the search process, helping administer the search and download the citation records to Rayyan (http://rayyan.qcri.org).

---Insert Table 1 here---

Inclusion and exclusion criteria

The inclusion criteria were developed *a priori* and listed in Table 2. Studies will be excluded articles if they: (1) did not report findings on human beings (e.g., studies on mice), (2) did not focus on full facial features (e.g., research on retina or lip-cleft), (3) did not conduct research in a medical context (e.g., in the context of criminology), and (4) did not report empirical findings (e.g., editorial or comment papers).

---Insert Table 2 here---

Risk of bias assessment

To ensure quality of included studies, a risk of bias assessment will be conducted independently by two reviewers, using the Cochrane Collaboration evaluation framework [45]. The framework has seven domains: (1) random sequence generation, (2) allocation concealment,

(3) blinding of participants and personnel, (4) blinding of outcome assessment, (5) incomplete outcome data, (6) selective reporting, and (7) any other source of bias. The risk of bias will be evaluated independently by two reviewers. Potential discrepancies regarding the risk of bias will be resolved via group discussions till a consensus is reached.

Data Extraction

Two reviewers will independently examine the citations and select studies for inclusion. Discrepancies will be resolved by group discussions till a consensus is reached. Data will be extracted based on research purpose and selection criteria adopted in this study. For articles that meet the inclusion criteria, the reviewers will extract the following information from the included papers: research purpose/questions, disease context, sample characteristics (e.g., characteristics of facial records), AI characteristics (e.g., algorithm adopted), and empirical findings.

Data synthesis and analysis

If eligible studies share enough similarities to be pooled, a meta-analysis will be conducted to gain further insights of the data. Main clinical, methodological, as well as statistical variances will be carefully considered to determine heterogeneity of the eligible studies. If eligible studies are found heterogeneous, a narrative synthesis will be conducted to summarize the data. A summary of the data extracted will be organized to synthesize key results. Both tables and graphs will be used to represent key characteristics of eligible articles. Descriptive analysis will be performed on categorical variables.

Discussion

Though a growing body of research has applied deep learning-based facial image analysis in the medical context for disease detection, diagnosis, and prognosis, to date, no systematic review has investigated the state-of-the-art application of deep learning-based facial image analysis recognition in addressing medical diagnoses and clinical states. Therefore, to bridge this gap, we aim to systematically review the literature and present the characteristics, challenges, as well as opportunities in deep learning-based facial analysis techniques applied in medical research. To better organize the research findings, we developed a framework that illustrates the main causes for abnormal facial expressions in patients. It is important to note that we are identifying medical states and conditions and not individuals.

After reviewing the literature [46-61], we identified the following four preliminary categories of causes for short-term or long-term abnormal facial expressions in people: (1) gene-related factors, (2) neurological factors, (3) psychiatric conditions, and (4) medication-induced triggers. Genetic-related factors, such as the presence or mutation of a certain gene, are the most studied cause for abnormal facial changes in individuals [35, 36, 40]. Down syndrome, which is affected by the presence of a third copy of chromosome 21, is an example of genetic-related factors that can cause individuals' abnormal facial changes [35]. Neurological factors can also cause individuals' facial phenotypes. Stroke or transient ischemic attack is an example of neurological factors, which can occur either prior to or after the onset of the disease [62, 63]. The third cause for abnormal facial changes centers on individuals' psychiatric conditions or mental illnesses, especially psychotic disorders such as the Tourette syndrome (facial tics). Last but not the least, medication-induced triggers, such as the Neuroleptic malignant syndrome (caused by antipsychotic medications), can also cause abnormal facial changes in people. Details of this framework can be found in Table 3. This framework will be used in the planned systematic review study to guide data extract process.

---Insert Table 3 here---

Overall, insights gained from this study will be able to provide much-needed timely understanding of the characteristics, challenges, as well as opportunities in the context of deep learning-based facial image analysis technologies applied in disease detection, diagnosis, and prognosis. In addition to gaining a connected and comprehensive understanding of current application of facial image analysis, results of the current study will also be able to shed light on whether, similar to facial recognition used in non-medical [64, 65] and medical contexts [66, 67], whether or to what degree is systematic bias is present in the application of deep learning technologies for facial image analysis. A biased and inaccurate facial image analysis system will not only exert unwarranted, though avoidable, disparities on patients (e.g., gender inequality) [66], it will also alienate the patients from the much-needed deep-learning-assisted medical opportunities their health and wellbeing can benefit from [68]. Therefore, for the consideration of patients' welfare and the development of the clinical practice, a timely understanding of the scope of the research literature as well as the challenges and opportunities faced by research on deep-learning-based facial image analysis is much needed.

Abbreviations

AI: Artificial intelligence

Ethics and dissemination

- Not applicable.

Authors' contributions

- ZS developed the research idea and drafted the manuscript. BL, FS, and JG, and JW reviewed and revised the manuscript.

Funding statement

- None.

Competing interests

- None.

Acknowledgements

- The authors wish to thank Emme Lopez, for her assistance in the search strategy development. Furthermore, the authors are very grateful for the constructive input offered by the editors and reviewers.

Word count:

- 2,228

Reference:

1. Wang, X., et al., *Artificial intelligence in tongue diagnosis: Using deep convolutional neural network for recognizing unhealthy tongue with tooth-mark.* Computational and structural biotechnology journal, 2020. **18**: p. 973-980.
2. Liang, B., et al., *Identification of the facial features of patients with cancer: A deep learning–based pilot study.* J Med Internet Res, 2020. **22**(4): p. e17234.
3. LeCun, Y., Y. Bengio, and G. Hinton, *Deep learning.* Nature, 2015. **521**(7553): p. 436-444.
4. Kamdar, M.M., et al., *A randomized controlled trial of a novel artificial-intelligence based smartphone application to optimize the management of cancer-related pain.* Journal of Clinical Oncology, 2018. **36**(34_suppl): p. 76-76.
5. Ström, P., et al., *Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: A population-based, diagnostic study.* The Lancet Oncology, 2020. **21**(2): p. 222-232.
6. Samuel, A.L., *Some studies in machine learning using the game of checkers.* IBM Journal of Research and Development, 1959. **3**(3): p. 210-229.
7. Komura, D. and S. Ishikawa, *Machine learning methods for histopathological image analysis.* Computational and structural biotechnology journal, 2018. **16**: p. 34-42.
8. Mitchell, T.M., *Machine Learning* International ed. 1997, New York, NY: McGraw-Hill.
9. Deng, L. and D. Yu, *Deep learning: Methods and applications.* Foundations and Trends in Signal Processing, 2014. **7**(3–4): p. 197–387.

10. Gilani, S.Z. and A. Mian. *Learning from millions of 3D scans for large-scale 3D face recognition*. in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018.
11. Algabri, R. and M.-T. Choi, *Deep-learning-based indoor human following of mobile robot using color feature.* Sensors, 2020. **20**(9).
12. Yin, J., et al., *Table cleaning task by human support robot using deep learning technique.* Sensors (Basel, Switzerland), 2020. **20**(6): p. 1698.
13. Balado, J., et al., *Road environment semantic segmentation with deep learning from MLS point cloud data.* Sensors, 2019. **19**(16): p. 3466.
14. Naqvi, R.A., et al., *Deep learning-based gaze detection system for automobile drivers using a NIR camera sensor.* Sensors (Basel, Switzerland), 2018. **18**(2): p. 456.
15. Liu, G., et al., *Automatic classification of esophageal lesions in endoscopic images using a convolutional neural network.* Annals of translational medicine, 2020. **8**(7): p. 486-486.
16. Chartrand, G., et al., *Deep learning: A primer for radiologists.* RadioGraphics, 2017. **37**(7): p. 2113-2131.
17. Jeong, Y., et al., *Deep learning-based decision support system for the diagnosis of neoplastic gallbladder polyps on ultrasonography: Preliminary results.* Scientific Reports, 2020. **10**(1): p. 7700.
18. Sánchez Fernández, I., et al., *Deep learning in rare disease. Detection of tubers in tuberous sclerosis complex.* PLoS ONE, 2020. **15**(4): p. e0232376-e0232376.
19. Marro, A., T. Bandukwala, and W. Mak, *Three-dimensional printing and medical imaging: A review of the methods and applications.* Current Problems in Diagnostic Radiology, 2016. **45**(1): p. 2-9.

20. Atutornu, J. and C.M. Hayre, *Personalised medicine and medical imaging: Opportunities and challenges for contemporary health care.* Journal of Medical Imaging and Radiation Sciences, 2018. **49**(4): p. 352-359.

21. Afshar, P., et al., *From handcrafted to deep-learning-based cancer radiomics: Challenges and opportunities.* IEEE Signal Processing Magazine, 2019. **36**(4): p. 132-160.

22. Antropova, N., B.Q. Huynh, and M.L. Giger, *A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets.* Medical Physics, 2017. **44**(10): p. 5162-5171.

23. Giger, M.L., *Machine learning in medical imaging.* Journal of the American College of Radiology, 2018. **15**(3, Part B): p. 512-520.

24. Sahiner, B., et al., *Deep learning in medical imaging and radiation therapy.* Medical Physics, 2019. **46**(1): p. e1-e36.

25. Huang, S., et al., *Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges.* Cancer Letters, 2020. **471**: p. 61-71.

26. Litjens, G., et al., *Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis.* Scientific Reports, 2016. **6**(1): p. 26286.

27. Yang, J., et al., *Development and validation of deep learning algorithms for scoliosis screening using back images.* Communications Biology, 2019. **2**(1): p. 390.

28. Moitra, D. and R.K. Mandal, *Prediction of non-small cell lung cancer histology by a deep ensemble of convolutional and bidirectional recurrent neural network.* Journal of Digital Imaging, 2020.

29. Mahbod, A., et al., *Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification.* Computer Methods and Programs in Biomedicine, 2020. **193**: p. 105475.

30. Li, Z., et al., *Chronic gastritis classification using gastric X-ray images with a semi-supervised learning method based on tri-training.* Medical & Biological Engineering & Computing, 2020. **58**(6): p. 1239-1250.

31. Liu, S., et al., *Isocitrate dehydrogenase (IDH) status prediction in histopathology images of gliomas using deep learning.* Scientific reports, 2020. **10**(1): p. 7733-7733.

32. Kim, Y.D., et al., *Effects of hypertension, diabetes, and smoking on age and sex prediction from retinal fundus images.* Scientific Reports, 2020. **10**(1): p. 4623.

33. Lam, C., et al., *Automated detection of diabetic retinopathy using deep learning.* AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science, 2018. **2017**: p. 147-155.

34. Martin, D., et al., *Systematic review and meta-analysis of the relationship between genetic risk for schizophrenia and facial emotion recognition.* Schizophrenia Research, 2020.

35. Kruszka, P., et al., *Down syndrome in diverse populations.* American Journal of Medical Genetics Part A, 2017. **173**(1): p. 42-53.

36. Latorre-Pellicer, A., et al., *Evaluating Face2Gene as a tool to identify Cornelia de Lange syndrome by facial phenotypes.* International Journal of Molecular Sciences, 2020. **21**(3).

37. Myers, L., et al., *Clinical versus automated assessments of morphological variants in twins with and without neurodevelopmental disorders.* American Journal of Medical Genetics Part A, 2020. **182**(5): p. 1177-1189.

38. Gurovich, Y., et al., *Identifying facial phenotypes of genetic disorders using deep learning.* Nature Medicine, 2019. **25**(1): p. 60-64.

39. Wang, Y. and M. Kosinski, *Deep neural networks are more accurate than humans at detecting sexual orientation from facial images.* Journal of Personality and Social Psychology, 2018. **114**(2): p. 246-257.

40. Kruszka, P., et al., *22q11.2 deletion syndrome in diverse populations.* American journal of medical genetics. Part A, 2017. **173**(4): p. 879-888.

41. Elmas, M. and B. Gogus, *Success of face analysis technology in rare genetic diseases diagnosed by whole-exome sequencing: A single-center experience.* Molecular Syndromology, 2020. **11**(1): p. 4-14.

42. Xu, C., et al., *Protocol registration or development may benefit the design, conduct and reporting of dose-response meta-analysis: Empirical evidence from a literature survey.* BMC Medical Research Methodology, 2019. **19**(1): p. 78.

43. dos Santos, M.B.F., et al., *Protocol registration improves reporting quality of systematic reviews in dentistry.* BMC Medical Research Methodology, 2020. **20**(1): p. 57.

44. Moher, D., et al., *Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement.* Systematic Reviews, 2015. **4**(1): p. 1-1.

45. Higgins, J.P., et al., *Cochrane handbook for systematic reviews of interventions*. 2019: John Wiley & Sons.

46. Kruszka, P., et al., *Down syndrome in diverse populations.* Am J Med Genet A, 2017. **173**(1): p. 42-53.

47. Basel-Vanagaite, L., et al., *Recognition of the Cornelia de Lange syndrome phenotype with facial dysmorphology novel analysis.* Clinical Genetics, 2016. **89**(5): p. 557-563.

48. Cha, S., et al., *Identification of five novel genetic loci related to facial morphology by genome-wide association studies.* BMC Genomics, 2018. **19**(1): p. 481.

49. Richmond, S., et al., *Facial genetics: A brief overview.* Frontiers in Genetics, 2018. **9**(462).

50. Spencer, C.R. and R.M. Irving, *Causes and management of facial nerve palsy.* British Journal of Hospital Medicine, 2016. **77**(12): p. 686-691.

51. Fasano, A. and M. Tinazzi, *Functional facial and tongue movement disorders*, in *Handbook of Clinical Neurology*, M. Hallett, J. Stone, and A. Carson, Editors. 2016, Elsevier. p. 353-365.

52. Hagberg, B., *Clinical manifestations and stages of rett syndrome.* Mental Retardation and Developmental Disabilities Research Reviews, 2002. **8**(2): p. 61-65.

53. Tan, N.C., L.L. Chan, and E.K. Tan, *Hemifacial spasm and involuntary facial movements.* QJM: An International Journal of Medicine, 2002. **95**(8): p. 493-500.

54. Liu, T.-L., et al., *Association between facial emotion recognition and bullying involvement among adolescents with high-functioning autism spectrum disorder.* International Journal of Environmental Research and Public Health, 2019. **16**(24): p. 5125.

55. Boutrus, M., et al., *Increased facial asymmetry in autism spectrum conditions is associated with symptom presentation.* Autism Research, 2019. **12**(12): p. 1774-1783.

56. Gill, C.E. and K. Kompoliti, *Clinical Features of Tourette Syndrome.* Journal of Child Neurology, 2019. **35**(2): p. 166-174.

57. Kim, D.D., et al., *Antipsychotic-associated symptoms of Tourette syndrome: A systematic review.* CNS Drugs, 2018. **32**(10): p. 917-938.

58. Thomas, C.S., *A study of facial dysmorphophobia.* Psychiatric Bulletin, 1995. **19**(12): p. 736-739.

59. Phillips, K.A., T.F. Cash, and L. Smolak, *Body image and body dysmorphic disorder.* Eating Disorders and Obesity: A Comprehensive Handbook, 2002. **2**: p. 115.

60. Reese, H.E., R.J. McNally, and S. Wilhelm, *Facial asymmetry detection in patients with body dysmorphic disorder.* Behaviour Research and Therapy, 2010. **48**(9): p. 936-940.

61. White, Janson J., et al., *DVL3 alleles resulting in a −1 frameshift of the last exon mediate autosomal-dominant Robinow syndrome.* The American Journal of Human Genetics, 2016. **98**(3): p. 553-561.

62. Onder, H., L. Albayrak, and H. Polat, *Frontal lobe ischemic stroke presenting with peripheral type facial palsy: A crucial diagnostic challenge in emergency practice.* Turkish Journal of Emergency Medicine, 2017. **17**(3): p. 112-114.

63. Han, Y.-y., et al., *Limb-shaking transient ischemic attack with facial muscles involuntary twitch successfully treated with internal carotid artery stenting.* Brain and Behavior, 2020. **n/a**(n/a): p. e01679.

64. Buolamwini, J. and T. Gebru, *Gender shades: Intersectional accuracy disparities in commercial gender classification*, in *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, A.F. Sorelle and W. Christo, Editors. 2018, PMLR: Proceedings of Machine Learning Research. p. 77--91.

65. Drozdowski, P., et al., *Demographic bias in biometrics: A survey on an emerging challenge.* IEEE Transactions on Technology and Society, 2020. **1**(2): p. 89-103.

66. Parikh, R.B., S. Teeple, and A.S. Navathe, *Addressing bias in artificial intelligence in health care.* JAMA, 2019. **322**(24): p. 2377-2378.

67. Obermeyer, Z., et al., *Dissecting racial bias in an algorithm used to manage the health of populations.* Science, 2019. **366**(6464): p. 447.

68. Garcia, R.V., et al. *The harms of demographic bias in feep face recognition research*. in *2019 International Conference on Biometrics (ICB)*. 2019.

Table 1. Example PubMed search strategy

Concept	**Search string**
Deep learning	"deep learning"[MeSH] OR "deep learning"[TIAB] OR "artificial intelligence" [MeSH] OR "artificial intelligence" [TIAB] OR "machine learning"[MeSH] OR "machine learning"[TIAB] OR "convolutional neural network"[MeSH] OR "convolutional neural network"[TIAB] OR "convolutional neural networks"[TIAB]
Facial image analysis	"face detect*" OR "facial detect*" OR "face recogn*" OR "facial recogn*" OR "face extract*" or "facial extract*" OR "face analys*" OR "facial analys*" OR "face dysmorphology" OR "facial dysmorphology" OR "face phenotype*" OR "facial phenotype*" OR "face feature*" OR "facial feature*" OR "face2gene" OR "gestalt theory" OR "face photograph*" OR "facial photograph*" OR "facial expression"

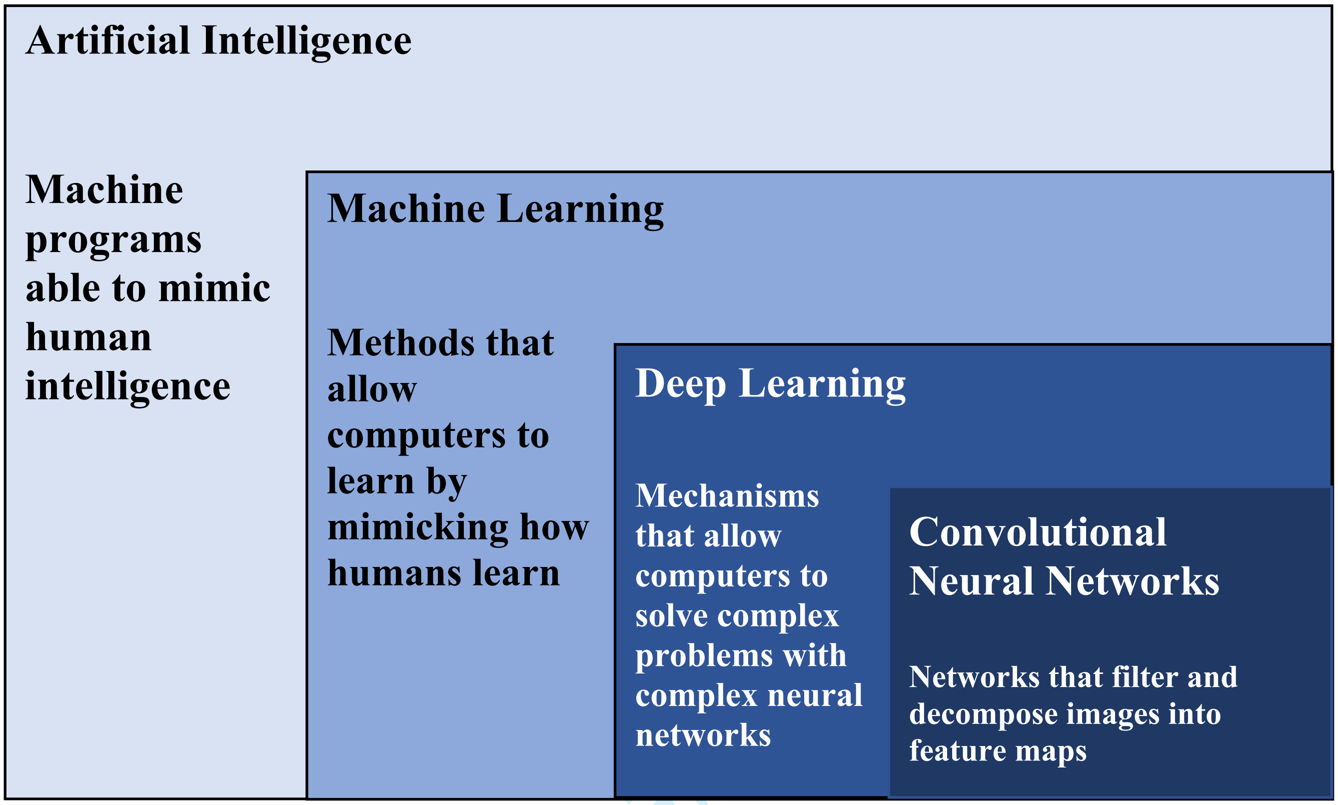
Table 2. Study inclusion criteria

Data type	**Inclusion criteria**
Participants	Individuals younger or older than 18 years old
Research context	Medical research or healthcare
Analytical technique	Deep learning algorithms-based facial image analysis
Language	English
Study type	Quantitative empirical study
Outcome	Report empirical and original findings on the application of deep learning-based facial image analysis in medical context (e.g., accuracy of facial image analysis in detecting Down syndrome)

Table 3. Main causes for abnormal facial expressions

Cause	Definition and Example
Gene-related factors	Gene-related factors are causes for individuals' abnormal facial changes that root in the presence or mutation of one or a set of genes. ***Examples:*** Down syndrome (genetic root: presence of a third copy of chromosome 21) or Cornelia de Lange syndrome (genetic root: NIPBL or *SMC1A*, *SMC3*, *RAD21* or *HDAC8*, *BRD4* and *ANKRD11* genes) [46-48, 61].
Neurological factors	Neurological factors are defined as reasons that are associated with individuals' congenital or acquired disorders of nerves and the nervous system. Neurological factors can either be related to genetic or non-genetic factors, caused by irregularity in nerves associated with the brain or the face. ***Examples:*** Neurological factors with genetic causes (e.g., Rett syndrome, *MECP2* gene; Cervical or Cranial dystonia, *GNAL* gene) and without (e.g., embouchure dystonia, Oromandibular dystonia) [52, 53]; due to nerves associated with the brain (e.g., stroke) or the face (Bell's palsy or facial paralysis, Hemifacial Spasm) [49-51].
Psychiatric conditions	Psychiatric conditions, especially psychotic disorders, have the potential to cause abnormal facial expressions among individuals. Psychiatric conditions could be broadly defined as mental illnesses, whereas psychotic disorder factors are causes to abnormal facial expressions that root in individuals' impaired sense of reality. ***Examples:*** Non-drug-related Tourette syndrome (facial tics) or Autism (facial expression limitation) [54-56].
Medication-induced triggers	Medication-induced triggers could be understood as causes to individuals' short-term or long-term abnormal facial changes due to their adverse reactions to a certain medication of a type of medications. ***Examples:*** Neuroleptic malignant syndrome (antipsychotic drugs), Tardive dyskinesia (antipsychotic medications), or drug-related Tourette syndrome [57-60].

Figure 1. Relationship between artificial intelligence, machine learning, deep learning, and convolutional neural networks.



PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	✓
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	✓
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	✓
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	✓
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	✓
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	✓
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	✓
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	✓
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	✓
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	✓
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	✓
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	✓
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	✓
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	✓

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	✓
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	✓
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	NA
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	NA
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	NA
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	NA
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	NA
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	NA
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	NA
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	NA
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	NA
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	NA
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	NA

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

BMJ Open

Deep Learning-Based Facial Image Analysis in Medical Research: A Systematic Review Protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-047549.R1
Article Type:	Protocol
Date Submitted by the Author:	05-Aug-2021
Complete List of Authors:	Su, Zhaohui; UT Health San Antonio, Center on Smart and Connected Health Technologies, Mays Cancer Center, School of Nursing Liang, Bin; Chinese Academy of Medical Sciences and Peking Union Medical College Shi, Feng; Shanghai United Imaging Intelligence Gelfond, J; University of Texas Health Science Center at San Antonio, Epidemiology and Biostatistics Šegalo, Sabina; University of Sarajevo Wang, Jing; Florida State University, College of Nursing Jia, Peng; University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC); International Initiative on Spatial Lifecourse Epidemiology (ISLE), Hao, Xiaoning; National Health Commission of the People's Republic of China
Primary Subject Heading:	Public health
Secondary Subject Heading:	Health informatics, Public health, Diagnostics, Evidence based practice, Medical management
Keywords:	PUBLIC HEALTH, Information technology < BIOTECHNOLOGY & BIOINFORMATICS, Health informatics < BIOTECHNOLOGY & BIOINFORMATICS, Telemedicine < BIOTECHNOLOGY & BIOINFORMATICS, BIOTECHNOLOGY & BIOINFORMATICS

SCHOLARONE™
Manuscripts

Deep Learning-Based Facial Image Analysis in Medical Research: A Systematic Review Protocol

Zhaohui Su,[1] Bin Liang,[2] Feng Shi,[3] Jonathan Gelfond,[4] Sabina Šegalo,[5] Jing Wang,[6] Peng Jia,[7] Xiaoning Hao[8*]

1. Ph.D., Center on Smart and Connected Health Technologies, Mays Cancer Center, School of Nursing, UT Health San Antonio, San Antonio, Texas, 78229, U.S.A.; szh@utexas.edu

2. Ph.D., M.D., Department of Radiation Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 10010, China; leangbin@gmail.com

3. Ph.D., Department of Research and Development, Shanghai United Imaging Intelligence Co., Ltd., Shanghai 200232, China; feng.shi@united-imaging.com

4. M.D., Ph.D., Associate Professor, Department of Population Health Sciences, UT Health San Antonio, San Antonio, Texas, 78229, U.S.A.; gelfondjal@uthscsa.edu

5. Ph.D., Department of Microbiology, Faculty of Medicine, University of Sarajevo, Sarajevo 71000, Bosnia & Herzegovina; sabina.segalo11@gmail.com

6. Ph.D., M.P.H., R.N., FAAN, Dean and Professor, College of Nursing, Florida State University, Tallahassee, Florida, 32306, U.S.A., jwang19@fsu.edu

7. Ph.D., Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China; International Institute of Spatial Lifecourse Epidemiology (ISLE), Hong Kong, China; jiapengff@hotmail.com

8. Ph.D., Professor, Director of Division, Division of Health Security Research, China National Health Development Research Center, National Health Commission, P.R. China; Beijing, 100044, China; haoxn@nhei.cn

*Corresponding author: Xiaoning Hao, Ph.D., Professor, Director of Division, Division of Health Security Research, China National Health Development Research Center, National Health Commission, P.R. China; No. 9 Chegongzhuang Street, Xicheng District, 100044, China; haoxn@nhei.cn; (+86) 010-88385748

Deep Learning-Based Facial Image Analysis in Medical Research: A Systematic Review Protocol

Abstract

Introduction: Deep learning techniques are gaining momentum in medical research. Evidence shows that deep learning has advantages over humans in image identification and classification, such as facial image analysis in detecting people's medical conditions. While positive findings are available, little is known about the state-of-the-art of deep learning-based facial image analysis in the medical context. For the consideration of patients' welfare and the development of the practice, a timely understanding of the challenges and opportunities faced by research on deep-learning-based facial image analysis is needed. To address this gap, we aim to conduct a systematic review to identify the characteristics and effects of deep learning-based facial image analysis in medical research. Insights gained from this systematic review will provide a much-needed understanding of the characteristics, challenges, as well as opportunities in deep learning-based facial image analysis applied in the contexts of disease detection, diagnosis, and prognosis.

Methods: Databases including PubMed, PsycINFO, CINAHL, IEEEXplore, and Scopus will be searched for relevant studies published in English in September, 2021. Titles, abstracts, and full-text articles will be screened to identify eligible articles. A manual search of the reference lists of the included articles will also be conducted. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework was adopted to guide the systematic review process. Two reviewers will independently examine the citations and select studies for inclusion. Discrepancies will be resolved by group discussions till a consensus is reached. Data will be extracted based on the research objective and selection criteria adopted in this study.

Ethics and dissemination: As the study is a protocol for a systematic review, ethical approval is not required. The study findings will be disseminated via peer-reviewed publications and conference presentations.

Study Protocol Registration: PROSPERO CRD42020196473

Keywords: facial image analysis; artificial intelligence; deep learning; convolutional neural network; abnormal facial expressions; facial analysis

Strengths and limitations of this study

- This systematic review protocol follows the Preferred Reporting Items for Systematic Review and Meta- Analysis Protocols guidelines.
- By examining the characteristics and effects of deep learning-based facial image analysis in medical research, this systematic review bridges the gap in the literature.
- This review is limited to evidence on the use and application of deep learning technologies in patients' facial image identification and classification.
- Non-English databases will not be searched, which might limit the representativeness of the results.

Deep Learning-Based Facial Image Analysis in Medical Research: A Systematic Review Protocol

Background

As disease manifestations often show in various places in the human body, such as Down Syndromes can change patients' facial features, researchers have been investigating whether analyzing appearance features can facilitate early disease detection and identification [1-5]. One promising field is deep learning-based facial analysis [6-8]. Deep learning represents a powerful range of artificial intelligence (AI) algorithm that allows computers to tackle complex problems via capitalizing on neural networks, such as convolutional neural networks (CNNs), that are rich in neurons, layers, and interconnectivity (see Figure 1) [9]. Simply put, deep learning is a mechanism that allows computers to solve complex problems by neural network architecture. This ability to develop complex network structures gives deep learning a distinctive advantage: it can automatically transform raw data input into meaningful features that enable pattern identification [10]. Deep learning technique has revolutionary potential in practical and research fields [11]. In practice, as deep learning effectively identifies objects, traffic signs, and faces, its adaptations have been widely applied in designing robots and self-driving cars [12-15]. Deep learning has also been widely adopted in biomedical and clinical research, particularly in the field of medical imaging [16-19].

---Insert Figure 1 here---

Medical conditions are often diagnosed by means of tests, such as biopsy and diagnostic imaging. An example list of diseases that have been analyzed by deep learning technologies could be found in Table 1. As diagnostic imaging is noninvasive and can facilitate personalized medicine, it is a preferred test option for patients and healthcare practitioners [20, 21]. This, in turn, has contributed to the exponential growth of medical imaging data and the increasing need for boosting medical image processing power to formulate diagnosis swiftly [21, 22]. Compared to traditional computer aided diagnosis for analyzing medical imaging, such as hand-crafted radiomics for tumor detection, deep learning methods are superior in their ability to process large quantities of medical images accurately and cost-effectively, without exerting a heavy workload on radiologists [23-27]. Evidence shows that deep learning-based medical image analysis was able to increase accuracy rates in various disease contexts, such as the identification of spinal disorder [1] and lung cancer histology [28], classification of skin lesion [29] and chronic gastritis [30], and the prediction of tumor-related genes [31] and vascular diseases [32].

---Insert Table 1 here---

Applying the deep learning technique to perform facial recognition and analysis tasks, researchers found that the technique yielded superior results in identifying and classifying faces of people with cancer from those without [6]. Similarly, examining facial phenotypes of people with genetic disorders, findings indicate that the technique was effective and was able to yield an optimal 91% top-10 accuracy [33]. Evidence further indicates that, for some tasks involving identifying and classifying facial images, deep learning techniques have often performed on par or better than human beings [5, 7, 10, 34-36]. Comparing clinical and deep learning evaluations of microdeletion syndrome facial phenotypes, researchers found that deep learning outperformed clinical evaluations in terms of sensitivity and specificity by 96% [35]. These findings combined suggest that deep learning-based facial analysis technology has great potential to address complex medical challenges prevalent in healthcare. However, there have not been any systematic review on the state-of-the-art applications of deep learning-based facial analysis in non-invasively evaluating medical conditions. Therefore, to bridge this gap, we aim to systematically review the literature and identify the characteristics and effects of deep learning-based facial analysis techniques applied in medical research.

Methods and analysis

This systematic review was registered with the International Prospective Register of Systematic Reviews database or PROSPERO (CRD42020196473) *a priori* to improve research rigor [37, 38]. The Principles of the Preferred Reporting Items for Meta-Analysis protocol (PRISMA) was adopted to guide this systematic review [39]. Our search strategy incorporated medical subject heading (MeSH) and keyword terms for the concept of deep learning and facial analysis. The search strategy was developed in consultation with an academic librarian, and subsequently will be deployed to target databases, including PubMed, PsycINFO, CINAHL, IEEEXplore, and Scopus (Table 2). The search will be initiated in September, 2021. Studies will be limited to journal articles published in English. We will adopt two additional search mechanisms to locate eligible articles: (1) a manual search of the reference list of the included articles will be performed, and (2) a reverse search of papers that cited articles included in the final review via Google Scholar. An academic librarian will facilitate the search process, helping administer the search and download the citation records to Rayyan (http://rayyan.qcri.org).

---Insert Table 2 here---

Inclusion and exclusion criteria

The inclusion criteria were developed *a priori* and listed in Table 3. Studies will be excluded articles if they 1) did not report findings on human beings (e.g., studies on mice), 2) did not focus on full facial features (e.g., research on retina or lip-cleft), 3) did not conduct research in a medical context (e.g., in the context of criminology), and 4) did not report empirical findings (e.g., editorial or comment papers).

---Insert Table 3 here---

Risk of bias assessment

To ensure quality of included studies, a risk of bias assessment will be conducted independently by two reviewers, using the Cochrane Collaboration evaluation framework [40]. The framework has seven domains: (1) random sequence generation, (2) allocation concealment, (3) blinding of participants and personnel, (4) blinding of outcome assessment, (5) incomplete outcome data, (6) selective reporting, and (7) any other source of bias. The risk of bias will be evaluated independently by two reviewers. Potential discrepancies regarding the risk of bias will be resolved via group discussions till a consensus is reached.

Data Extraction

Two reviewers will independently examine the citations and select studies for inclusion. Discrepancies will be resolved by group discussions till a consensus is reached. Data will be extracted based on the research objective and selection criteria adopted in this study. For articles that meet the inclusion criteria, the reviewers will extract the following information from the included papers: research objective /questions, disease context, sample characteristics (e.g., characteristics of facial records), AI characteristics (e.g., algorithm adopted), and empirical findings.

Data synthesis and analysis

If eligible studies share enough similarities to be pooled, a meta-analysis will be conducted to gain further insights into the data. Main clinical, methodological, as well as statistical differences will be carefully considered to determine the heterogeneity of the eligible studies. If eligible studies are found heterogeneous, a narrative synthesis will be conducted to summarize the data. A summary of the data extracted will be organized to synthesize key results. Both tables and graphs will be used to represent the key characteristics of eligible articles. Descriptive analysis will be performed on categorical variables. In this review, we will undertake a narrative approach to synthesize data. In other words, in addition to shedding light on key information like the sensitivity, specificity, overall accuracy of the deep learning technologies in analyzing facial images (as opposed to clinicians' analyses), we will also provide detailed analysis of the disease contexts and the techniques applied to chart the state-of-the-art of deep learning technologies in facial image analyses.

Ethics and dissemination

As the study is a protocol for a systematic review, ethical approval is not required. The study findings will be disseminated via peer-reviewed publications and conference presentations.

Patient and Public Involvement

The nature of the study, which is a review and analysis of previously published data, dictates that there is limited to no meaningful need for patient and public involvement in the design, delivery, or dissemination of the research findings.

Discussion

Though a growing body of research has applied deep learning-based facial image analysis in the medical context for disease detection, diagnosis, and prognosis, to date, no systematic review has investigated the state-of-the-art application of deep learning-based facial image analysis recognition in addressing medical diagnoses and clinical states. Therefore, to bridge this gap, we aim to systematically review the literature and present the characteristics, challenges, as well as opportunities in deep learning-based facial analysis techniques applied in medical research. To better organize the research findings, we developed a framework that illustrates the main causes for abnormal facial expressions in patients. It is important to note that we are identifying medical states and conditions and not individuals.

After reviewing the literature [41-56], we identified the following four preliminary categories of causes for short-term or long-term abnormal facial expressions in people: (1) gene-related factors, (2) neurological factors, (3) psychiatric conditions, and (4) medication-induced triggers. Genetic-related factors, such as the presence or mutation of a certain gene, are the most studied cause for abnormal facial changes in individuals [5, 7, 35]. Down syndrome, which is affected by the presence of a third copy of chromosome 21, is an example of genetic-related factors that can cause individuals' abnormal facial changes [5]. Neurological factors can also cause individuals' facial phenotypes. Stroke or transient ischemic attack is an example of neurological factors, which can occur either prior to or after the onset of the disease [57, 58]. The third cause for abnormal facial changes centers on individuals' psychiatric conditions or mental illnesses, especially psychotic disorders such as the Tourette syndrome (facial tics). Last but not the least, medication-induced triggers, such as the Neuroleptic malignant syndrome (caused by antipsychotic medications), can also cause abnormal facial changes in people. Details of this framework can be found in Table 4. This framework will be used in the planned systematic review study to guide the data extraction process.

---Insert Table 4 here---

Overall, insights gained from this study will be able to provide a much-needed understanding of the characteristics, challenges, as well as opportunities in the context of deep learning-based facial image analysis technologies applied in disease detection, diagnosis, and prognosis. In addition to gaining a connected and comprehensive understanding of the current application of facial image analysis, results of the study will also be able to shed light on whether, similar to facial recognition used in non-medical [59, 60] and medical contexts [61, 62], whether or to what degree is systematic bias is present in the application of deep learning technologies for facial image analysis. A biased and inaccurate facial image analysis system will not only exert unwarranted, though avoidable, disparities on patients (e.g., gender inequality) [61], it will also alienate the patients from the much-needed deep-learning-assisted medical opportunities their health and wellbeing can benefit from [63]. Therefore, for the consideration of patients' welfare and the development of the clinical practice, a timely understanding of the scope of the research literature as well as the challenges and opportunities faced by research on deep-learning-based facial image analysis is much needed.

Abbreviations

AI: Artificial intelligence

Ethics and dissemination

- Not applicable.

Authors' contributions

- ZS developed the research idea and drafted the manuscript. BL, FS, JG, SS, JW, PJ, & XH reviewed and revised the manuscript.

Funding statement

- None.

Competing interests

- None.

Acknowledgements

- The authors wish to thank Emme Lopez, for her assistance in the search strategy development. Furthermore, the authors are very grateful for the constructive input offered by the editors and reviewers.

References:

1. Yang, J., et al., *Development and validation of deep learning algorithms for scoliosis screening using back images.* Communications Biology, 2019. **2**(1): p. 390.
2. Wang, X., et al., *Artificial intelligence in tongue diagnosis: Using deep convolutional neural network for recognizing unhealthy tongue with tooth-mark.* Computational and structural biotechnology journal, 2020. **18**: p. 973-980.
3. Lam, C., et al., *Automated detection of diabetic retinopathy using deep learning.* AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science, 2018. **2017**: p. 147-155.
4. Martin, D., et al., *Systematic review and meta-analysis of the relationship between genetic risk for schizophrenia and facial emotion recognition.* Schizophrenia Research, 2020.
5. Kruszka, P., et al., *Down syndrome in diverse populations.* American Journal of Medical Genetics Part A, 2017. **173**(1): p. 42-53.
6. Liang, B., et al., *Identification of the facial features of patients with cancer: A deep learning–based pilot study.* J Med Internet Res, 2020. **22**(4): p. e17234.
7. Latorre-Pellicer, A., et al., *Evaluating Face2Gene as a tool to identify Cornelia de Lange syndrome by facial phenotypes.* International Journal of Molecular Sciences, 2020. **21**(3).
8. Myers, L., et al., *Clinical versus automated assessments of morphological variants in twins with and without neurodevelopmental disorders.* American Journal of Medical Genetics Part A, 2020. **182**(5): p. 1177-1189.
9. Deng, L. and D. Yu, *Deep learning: Methods and applications.* Foundations and Trends in Signal Processing, 2014. **7**(3–4): p. 197–387.
10. LeCun, Y., Y. Bengio, and G. Hinton, *Deep learning.* Nature, 2015. **521**(7553): p. 436-444.
11. Gilani, S.Z. and A. Mian. *Learning from millions of 3D scans for large-scale 3D face recognition*. in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018.
12. Algabri, R. and M.-T. Choi, *Deep-learning-based indoor human following of mobile robot using color feature.* Sensors, 2020. **20**(9).
13. Yin, J., et al., *Table cleaning task by human support robot using deep learning technique.* Sensors (Basel, Switzerland), 2020. **20**(6): p. 1698.
14. Balado, J., et al., *Road environment semantic segmentation with deep learning from MLS point cloud data.* Sensors, 2019. **19**(16): p. 3466.
15. Naqvi, R.A., et al., *Deep learning-based gaze detection system for automobile drivers using a NIR camera sensor.* Sensors (Basel, Switzerland), 2018. **18**(2): p. 456.
16. Liu, G., et al., *Automatic classification of esophageal lesions in endoscopic images using a convolutional neural network.* Annals of translational medicine, 2020. **8**(7): p. 486-486.
17. Chartrand, G., et al., *Deep learning: A primer for radiologists.* RadioGraphics, 2017. **37**(7): p. 2113-2131.
18. Jeong, Y., et al., *Deep learning-based decision support system for the diagnosis of neoplastic gallbladder polyps on ultrasonography: Preliminary results.* Scientific Reports, 2020. **10**(1): p. 7700.
19. Sánchez Fernández, I., et al., *Deep learning in rare disease. Detection of tubers in tuberous sclerosis complex.* PLoS ONE, 2020. **15**(4): p. e0232376-e0232376.

20. Marro, A., T. Bandukwala, and W. Mak, *Three-dimensional printing and medical imaging: A review of the methods and applications.* Current Problems in Diagnostic Radiology, 2016. **45**(1): p. 2-9.
21. Atutornu, J. and C.M. Hayre, *Personalised medicine and medical imaging: Opportunities and challenges for contemporary health care.* Journal of Medical Imaging and Radiation Sciences, 2018. **49**(4): p. 352-359.
22. Afshar, P., et al., *From handcrafted to deep-learning-based cancer radiomics: Challenges and opportunities.* IEEE Signal Processing Magazine, 2019. **36**(4): p. 132-160.
23. Antropova, N., B.Q. Huynh, and M.L. Giger, *A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets.* Medical Physics, 2017. **44**(10): p. 5162-5171.
24. Giger, M.L., *Machine learning in medical imaging.* Journal of the American College of Radiology, 2018. **15**(3, Part B): p. 512-520.
25. Sahiner, B., et al., *Deep learning in medical imaging and radiation therapy.* Medical Physics, 2019. **46**(1): p. e1-e36.
26. Huang, S., et al., *Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges.* Cancer Letters, 2020. **471**: p. 61-71.
27. Litjens, G., et al., *Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis.* Scientific Reports, 2016. **6**(1): p. 26286.
28. Moitra, D. and R.K. Mandal, *Prediction of non-small cell lung cancer histology by a deep ensemble of convolutional and bidirectional recurrent neural network.* Journal of Digital Imaging, 2020.
29. Mahbod, A., et al., *Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification.* Computer Methods and Programs in Biomedicine, 2020. **193**: p. 105475.
30. Li, Z., et al., *Chronic gastritis classification using gastric X-ray images with a semi-supervised learning method based on tri-training.* Medical & Biological Engineering & Computing, 2020. **58**(6): p. 1239-1250.
31. Liu, S., et al., *Isocitrate dehydrogenase (IDH) status prediction in histopathology images of gliomas using deep learning.* Scientific reports, 2020. **10**(1): p. 7733-7733.
32. Kim, Y.D., et al., *Effects of hypertension, diabetes, and smoking on age and sex prediction from retinal fundus images.* Scientific Reports, 2020. **10**(1): p. 4623.
33. Gurovich, Y., et al., *Identifying facial phenotypes of genetic disorders using deep learning.* Nature Medicine, 2019. **25**(1): p. 60-64.
34. Wang, Y. and M. Kosinski, *Deep neural networks are more accurate than humans at detecting sexual orientation from facial images.* Journal of Personality and Social Psychology, 2018. **114**(2): p. 246-257.
35. Kruszka, P., et al., *22q11.2 deletion syndrome in diverse populations.* American journal of medical genetics. Part A, 2017. **173**(4): p. 879-888.
36. Elmas, M. and B. Gogus, *Success of face analysis technology in rare genetic diseases diagnosed by whole-exome sequencing: A single-center experience.* Molecular Syndromology, 2020. **11**(1): p. 4-14.
37. Xu, C., et al., *Protocol registration or development may benefit the design, conduct and reporting of dose-response meta-analysis: Empirical evidence from a literature survey.* BMC Medical Research Methodology, 2019. **19**(1): p. 78.

38. dos Santos, M.B.F., et al., *Protocol registration improves reporting quality of systematic reviews in dentistry.* BMC Medical Research Methodology, 2020. **20**(1): p. 57.
39. Moher, D., et al., *Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement.* Systematic Reviews, 2015. **4**(1): p. 1-1.
40. Higgins, J.P., et al., *Cochrane handbook for systematic reviews of interventions*. 2019: John Wiley & Sons.
41. Kruszka, P., et al., *Down syndrome in diverse populations.* Am J Med Genet A, 2017. **173**(1): p. 42-53.
42. Basel-Vanagaite, L., et al., *Recognition of the Cornelia de Lange syndrome phenotype with facial dysmorphology novel analysis.* Clinical Genetics, 2016. **89**(5): p. 557-563.
43. Cha, S., et al., *Identification of five novel genetic loci related to facial morphology by genome-wide association studies.* BMC Genomics, 2018. **19**(1): p. 481.
44. Richmond, S., et al., *Facial genetics: A brief overview.* Frontiers in Genetics, 2018. **9**(462).
45. Spencer, C.R. and R.M. Irving, *Causes and management of facial nerve palsy.* British Journal of Hospital Medicine, 2016. **77**(12): p. 686-691.
46. Fasano, A. and M. Tinazzi, *Functional facial and tongue movement disorders*, in *Handbook of Clinical Neurology*, M. Hallett, J. Stone, and A. Carson, Editors. 2016, Elsevier. p. 353-365.
47. Hagberg, B., *Clinical manifestations and stages of rett syndrome.* Mental Retardation and Developmental Disabilities Research Reviews, 2002. **8**(2): p. 61-65.
48. Tan, N.C., L.L. Chan, and E.K. Tan, *Hemifacial spasm and involuntary facial movements.* QJM: An International Journal of Medicine, 2002. **95**(8): p. 493-500.
49. Liu, T.-L., et al., *Association between facial emotion recognition and bullying involvement among adolescents with high-functioning autism spectrum disorder.* International Journal of Environmental Research and Public Health, 2019. **16**(24): p. 5125.
50. Boutrus, M., et al., *Increased facial asymmetry in autism spectrum conditions is associated with symptom presentation.* Autism Research, 2019. **12**(12): p. 1774-1783.
51. Gill, C.E. and K. Kompoliti, *Clinical Features of Tourette Syndrome.* Journal of Child Neurology, 2019. **35**(2): p. 166-174.
52. Kim, D.D., et al., *Antipsychotic-associated symptoms of Tourette syndrome: A systematic review.* CNS Drugs, 2018. **32**(10): p. 917-938.
53. Thomas, C.S., *A study of facial dysmorphophobia.* Psychiatric Bulletin, 1995. **19**(12): p. 736-739.
54. Phillips, K.A., T.F. Cash, and L. Smolak, *Body image and body dysmorphic disorder.* Eating Disorders and Obesity: A Comprehensive Handbook, 2002. **2**: p. 115.
55. Reese, H.E., R.J. McNally, and S. Wilhelm, *Facial asymmetry detection in patients with body dysmorphic disorder.* Behaviour Research and Therapy, 2010. **48**(9): p. 936-940.
56. White, Janson J., et al., *DVL3 alleles resulting in a −1 frameshift of the last exon mediate autosomal-dominant Robinow syndrome.* The American Journal of Human Genetics, 2016. **98**(3): p. 553-561.
57. Onder, H., L. Albayrak, and H. Polat, *Frontal lobe ischemic stroke presenting with peripheral type facial palsy: A crucial diagnostic challenge in emergency practice.* Turkish Journal of Emergency Medicine, 2017. **17**(3): p. 112-114.

58. Han, Y.-y., et al., *Limb-shaking transient ischemic attack with facial muscles involuntary twitch successfully treated with internal carotid artery stenting.* Brain and Behavior, 2020. **n/a**(n/a): p. e01679.
59. Buolamwini, J. and T. Gebru, *Gender shades: Intersectional accuracy disparities in commercial gender classification*, in *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, A.F. Sorelle and W. Christo, Editors. 2018, PMLR: Proceedings of Machine Learning Research. p. 77--91.
60. Drozdowski, P., et al., *Demographic bias in biometrics: A survey on an emerging challenge.* IEEE Transactions on Technology and Society, 2020. **1**(2): p. 89-103.
61. Parikh, R.B., S. Teeple, and A.S. Navathe, *Addressing bias in artificial intelligence in health care.* JAMA, 2019. **322**(24): p. 2377-2378.
62. Obermeyer, Z., et al., *Dissecting racial bias in an algorithm used to manage the health of populations.* Science, 2019. **366**(6464): p. 447.
63. Garcia, R.V., et al. *The harms of demographic bias in feep face recognition research*. in *2019 International Conference on Biometrics (ICB)*. 2019.
64. Kong, X., et al., *Automatic detection of acromegaly from facial photographs using machine learning methods.* EBioMedicine, 2018. **27**: p. 94-102.
65. Liang, B., et al., *Identification of the facial features of patients with cancer: A deep learning-based pilot study.* J Med Internet Res, 2020. **22**(4): p. e17234.
66. Latorre-Pellicer, A., et al., *Evaluating Face2Gene as a tool to identify Cornelia de Lange syndrome by facial phenotypes.* Int J Mol Sci, 2020. **21**(3).
67. Lin, S., et al., *Feasibility of using deep learning to detect coronary artery disease based on facial photo.* Eur Heart J, 2020. **41**(46): p. 4400-4411.
68. Zhao, Q., et al., *Digital facial dysmorphology for genetic screening: Hierarchical constrained local model using ICA.* Med Image Anal, 2014. **18**(5): p. 699-710.
69. Goceri, E., *Deep learning based classification of facial dermatological disorders.* Computers in Biology and Medicine, 2021. **128**: p. 104118.
70. Han, S.S., et al., *Keratinocytic skin cancer detection on the face using region-based convolutional neural network.* JAMA Dermatology, 2020. **156**(1): p. 29-37.
71. Camino, A., et al., *Deep learning for the segmentation of preserved photoreceptors on en face optical coherence tomography in two inherited retinal diseases.* Biomed Opt Express, 2018. **9**(7): p. 3092-3105.
72. Bargshady, G., et al., *Enhanced deep learning algorithm development to detect pain intensity from facial expression images.* Expert Systems with Applications, 2020. **149**: p. 113305.
73. Yolcu, G., et al. *Deep learning-based facial expression recognition for monitoring neurological disorders*. in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. 2017.

Table 1. An example list of diseases that have been analyzed by deep learning techniques

Disease context	Deep Learning Technique
Acromegaly	Convolutional Neural Network (along with Generalized Linear Models; K-nearest neighbors; Support Vector Machines; forests of randomized trees) [64]
Cancer	Convolutional neutral network [65]
Cornelia de Lange syndrome	DeepGestalt technology [66]
Coronary artery disease	Convolutional neural network [67]
Down syndrome	Independent component analysis [68]
Facial dermatological disorders	Convolutional neural network [69]
Keratinocytic Skin Cancer	Convolutional neutral network [70]
Inherited retinal degenerations	Convolutional neural network [71]
Noonan syndrome	DeepGestalt technology [33]
Pain intensity	Convolutional neutral network [72]
Neurological disorders	Convolutional neutral network [73]

Table 2. Example PubMed search strategy

Concept	**Search string**
Deep learning	"deep learning"[MeSH] OR "deep learning"[TIAB] OR "artificial intelligence" [MeSH] OR "artificial intelligence" [TIAB] OR "machine learning"[MeSH] OR "machine learning"[TIAB] OR "convolutional neural network"[MeSH] OR "convolutional neural network"[TIAB] OR "convolutional neural networks"[TIAB]
Facial image analysis	"face detect*" OR "facial detect*" OR "face recogn*" OR "facial recogn*" OR "face extract*" or "facial extract*" OR "face analys*" OR "facial analys*" OR "face dysmorphology" OR "facial dysmorphology" OR "face phenotype*" OR "facial phenotype*" OR "face feature*" OR "facial feature*" OR "face2gene" OR "gestalt theory" OR "face photograph*" OR "facial photograph*" OR "facial expression"

Table 3. Study inclusion criteria

Data type	**Inclusion criteria**
Participants	Individuals younger or older than 18 years old
Research context	Medical research or healthcare
Analytical technique	Deep learning algorithms-based facial image analysis
Language	English
Study type	Quantitative empirical study
Outcome	Report empirical and original findings on the application of deep learning-based facial image analysis in medical context (e.g., accuracy of facial image analysis in detecting Down syndrome)

Table 4. Main causes for abnormal facial expressions

Cause	Definition and Example
Gene-related factors	Gene-related factors are causes for individuals' abnormal facial changes that root in the presence or mutation of one or a set of genes. ***Examples**:* Down syndrome (genetic root: presence of a third copy of chromosome 21) or Cornelia de Lange syndrome (genetic root: NIPBL or *SMC1A*, *SMC3*, *RAD21* or *HDAC8*, *BRD4* and *ANKRD11* genes) [41-43, 56].
Neurological factors	Neurological factors are defined as reasons that are associated with individuals' congenital or acquired disorders of nerves and the nervous system. Neurological factors can either be related to genetic or non-genetic factors, caused by irregularity in nerves associated with the brain or the face. ***Examples:*** Neurological factors with genetic causes (e.g., Rett syndrome, *MECP2* gene; Cervical or Cranial dystonia, *GNAL* gene) and without (e.g., embouchure dystonia, Oromandibular dystonia) [47, 48]; due to nerves associated with the brain (e.g., stroke) or the face (Bell's palsy or facial paralysis, Hemifacial Spasm) [44-46].
Psychiatric conditions	Psychiatric conditions, especially psychotic disorders, have the potential to cause abnormal facial expressions among individuals. Psychiatric conditions could be broadly defined as mental illnesses, whereas psychotic disorder factors are causes to abnormal facial expressions that root in individuals' impaired sense of reality. ***Examples:*** Non-drug-related Tourette syndrome (facial tics) or Autism (facial expression limitation) [49-51].
Medication-induced triggers	Medication-induced triggers could be understood as causes to individuals' short-term or long-term abnormal facial changes due to their adverse reactions to a certain medication of a type of medications. ***Examples:*** Neuroleptic malignant syndrome (antipsychotic drugs), Tardive dyskinesia (antipsychotic medications), or drug-related Tourette syndrome [52-55].

Figure 1. Relationship between artificial intelligence, machine learning, deep learning, and convolutional neural networks.

Please refer to the supplementary materials section for the figure

Figure 1. Relationship between artificial intelligence, machine learning, deep learning, and convolutional neural networks.

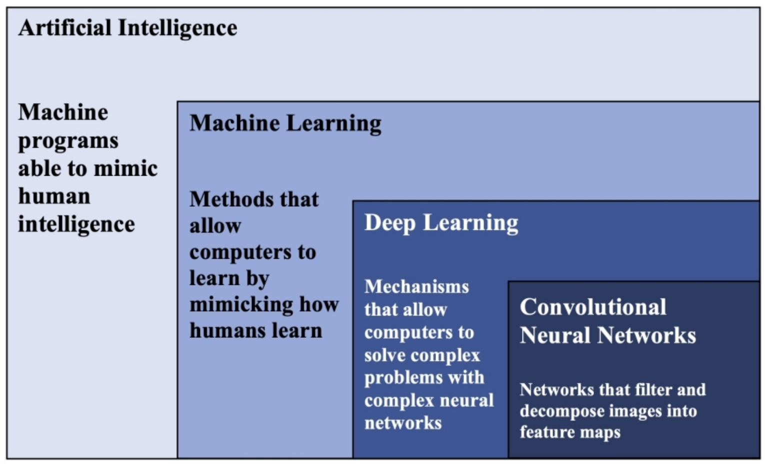


Figure 1. Relationship between artificial intelligence, machine learning, deep learning, and convolutional neural networks.

899x1164mm (72 x 72 DPI)