

The role of mobile health applications in improving patient engagement and health outcomes: A critical review

Chioma Anthonia Okolo ^{1,*}, Oloruntoba Babawarun ², Jeremiah Olawumi Awoogun ³, Adekunle Oyeyemi Adeniyi ⁴ and Rawlings Chidi ⁵

¹ Federal Medical Centre, Asaba, Delta State, Nigeria.

² Global Future Redemption Empowerment Foundation, Nigeria.

³ Bharat Serums and Vaccines Limited Lagos, Nigeria.

⁴ United Nations Population Fund, Sri Lanka.

⁵ Parkville MO & North Kansas City Hospital, US State.

International Journal of Science and Research Archive, 2024, 11(01), 2566–2574

Publication history: Received on 13 January 2024; revised on 19 February 2024; accepted on 22 February 2024

Article DOI: <https://doi.org/10.30574/ijsra.2024.11.1.0334>

Abstract

Mobile health applications (apps) have revolutionized healthcare delivery by enhancing patient engagement and improving health outcomes. This critical review examines the role of mobile health apps in improving patient engagement and health outcomes. The review discusses the benefits of mobile health apps in empowering patients to manage their health, facilitating communication between patients and healthcare providers, and promoting healthy behaviors. Mobile health apps have transformed the way patients engage with their healthcare by providing them with convenient access to health information and resources. These apps enable patients to track their health metrics, such as physical activity, diet, and medication adherence, allowing for better self-management of chronic conditions. By empowering patients to take control of their health, mobile health apps can lead to improved health outcomes and reduced healthcare costs. Furthermore, mobile health apps facilitate communication between patients and healthcare providers, enabling remote monitoring and virtual consultations. This improved communication can lead to more timely interventions and better coordination of care, ultimately improving patient outcomes. Mobile health apps also promote healthy behaviors by providing patients with personalized recommendations and reminders to engage in preventive care activities. Despite their potential benefits, mobile health apps also face challenges, such as ensuring data privacy and security, and addressing disparities in access to technology. Future research should focus on evaluating the effectiveness of mobile health apps in improving patient engagement and health outcomes, and on developing strategies to overcome these challenges. In conclusion, mobile health apps have the potential to significantly improve patient engagement and health outcomes. By empowering patients to manage their health, facilitating communication with healthcare providers, and promoting healthy behaviors, mobile health apps are transforming healthcare delivery and improving patient care.

Keywords: Role; Mobile health; Application; Patient engagement; Health outcomes

1. Introduction

Mobile health applications (apps) have emerged as powerful tools in modern healthcare, offering innovative ways to improve patient engagement and enhance health outcomes. These apps, designed for use on smartphones and other mobile devices, provide users with access to a wide range of health-related information, resources, and services (Iribarren, et. al. 2021, Nguyen, et. al., 2019, Wei, et. al., 2020).

* Corresponding author: Chioma Anthonia Okolo

Patient engagement is a critical component of healthcare delivery, as it empowers patients to take an active role in managing their health. Engaged patients are more likely to adhere to treatment plans, adopt healthy behaviors, and achieve better health outcomes. Mobile health apps play a key role in promoting patient engagement by providing users with tools to track their health metrics, access personalized health information, and communicate with healthcare providers (Schneider-Kamp & Askegaard, 2020, Tanniru, 2019, Tobiano, Jeroke-Owen & Marshall, 2021).

The purpose of this review is to critically evaluate the role of mobile health applications in improving patient engagement and health outcomes. The review will examine the benefits of mobile health apps in empowering patients, facilitating communication between patients and healthcare providers, and promoting healthy behaviors. It will also discuss the challenges and limitations associated with mobile health apps, as well as strategies for enhancing their effectiveness. By providing a comprehensive overview of the current state of mobile health applications in healthcare, this review aims to inform healthcare professionals, researchers, and policymakers about the potential benefits and challenges of integrating mobile health apps into healthcare delivery.

2. Mobile Health Applications and Patient Engagement

Mobile health applications (apps) have revolutionized the way patients engage with their healthcare by providing convenient access to health information, tools for self-management, and communication channels with healthcare providers. Patient engagement, defined as the involvement of patients in their healthcare decisions and treatment plans, is crucial for improving health outcomes and overall healthcare quality. This review explores the role of mobile health apps in promoting patient engagement, highlighting their benefits and providing examples of apps designed to engage patients (Ghose, et. al., 2021, Song, 2021, Vaghefi & Tulu, 2019).

Patient engagement is a multidimensional concept that refers to the active involvement of patients in their healthcare journey. It encompasses a range of behaviors, including seeking health information, participating in treatment decisions, adhering to treatment plans, and managing chronic conditions. Engaged patients are more likely to achieve positive health outcomes, experience higher satisfaction with their care, and have lower healthcare costs. Mobile health apps play a crucial role in promoting patient engagement by providing patients with tools and resources to actively manage their health (Cengiz & Korkmaz, 2022, Graffigna, et. al., 2020). These apps empower patients to track their health metrics, monitor their progress, and communicate with healthcare providers. By facilitating access to personalized health information and resources, mobile health apps help patients make informed decisions about their health and treatment options.

MyFitnessPal is a popular app that helps users track their diet, exercise, and weight loss goals. It provides personalized recommendations based on users' health goals and allows them to track their progress over time. Mango Health is an app that helps users manage their medications by providing reminders to take their medications, tracking their medication adherence, and providing information about potential drug interactions (Bracken & Waite, 2020, Gordon, Althoff & Leskovec, 2019, Patel, et. al., 2019). Fitbit is a wearable device and app that tracks users' physical activity, sleep patterns, and heart rate. It provides users with personalized insights and recommendations to help them improve their overall health and well-being. Ada is an AI-powered app that helps users assess their symptoms and provides personalized health recommendations. It allows users to track their symptoms over time and share this information with their healthcare providers.

In conclusion, mobile health apps play a crucial role in promoting patient engagement by providing patients with tools and resources to actively manage their health. By empowering patients to take control of their health, these apps help improve health outcomes and enhance the overall healthcare experience.

3. Brief History of Mobile Health Applications

Mobile health applications (apps) have a rich history of improving patient engagement and health outcomes, evolving alongside advancements in mobile technology and healthcare practices. Here is a brief overview of the key milestones in the history of mobile health apps: The early 2000s saw the emergence of the first mobile health apps, primarily focused on simple tasks such as medication reminders and symptom tracking. These apps were basic in functionality and often lacked the sophistication seen in later developments. The late 2000s marked the rise of smartphones, which revolutionized the mobile health app landscape. The introduction of app stores, such as the Apple App Store and Google Play Store, provided a platform for developers to create more advanced health apps. The 2010s saw a rapid expansion of health and fitness apps, catering to a wide range of needs such as diet tracking, exercise monitoring, and sleep tracking.

(Galetsi, Katsaliaki & Kumar, 2023, Nussbaum, et. Al., 2019). These apps were instrumental in promoting healthy behaviors and encouraging users to take a proactive approach to their health.

The mid-2010s saw the integration of mobile health apps with wearable devices, such as fitness trackers and smartwatches. This integration allowed for more comprehensive health monitoring and provided users with real-time feedback on their health metrics. In the late 2010s, there was a growing focus on using mobile health apps for chronic disease management. Apps were developed to help users manage conditions such as diabetes, hypertension, and asthma, providing personalized care plans and monitoring tools. Today, mobile health apps are incorporating artificial intelligence (AI) and machine learning (ML) algorithms to provide more personalized and accurate health recommendations (Azizan, Ahmed & Razak, 2023, Blaszka & Rascon, 2023, Waalen, 2023). These apps can analyze user data to provide tailored advice and identify early signs of health problems.

Overall, the history of mobile health applications demonstrates a continuous evolution towards more sophisticated and effective tools for improving patient engagement and health outcomes. As technology continues to advance, mobile health apps are poised to play an even greater role in shaping the future of healthcare delivery.

4. Impact of Mobile Health Applications on Health Outcomes

Mobile health applications (apps) have become increasingly popular tools for improving health outcomes by providing users with access to health information, monitoring tools, and behavior change support (Debon, et. al., 2019, Iribarren, et. al., 2021, Mahmood, et. al., 2019). This review explores the impact of mobile health apps on health outcomes, including the evidence supporting their effectiveness, studies demonstrating improved outcomes, and factors influencing their impact. Numerous studies have demonstrated the effectiveness of mobile health apps in improving health outcomes across various populations and health conditions. These apps have been shown to improve medication adherence, manage chronic conditions, promote healthy behaviors, and enhance patient-provider communication. For example, a study published in JAMA Internal Medicine found that patients using a mobile app to track their blood pressure had lower blood pressure readings compared to those who did not use the app. Another study published in the Journal of Medical Internet Research showed that patients with diabetes who used a mobile app to track their blood glucose levels experienced improved glycemic control compared to those who did not use the app.

Several studies have demonstrated the positive impact of mobile health apps on health outcomes. For instance, a study published in Diabetes Care found that patients with type 2 diabetes who used a mobile app to track their diet and exercise habits experienced greater weight loss and improvements in their glycemic control compared to those who did not use the app. Similarly, a study published in the American Journal of Preventive Medicine showed that patients who used a mobile app to track their physical activity levels increased their daily step count and improved their cardiovascular fitness compared to those who did not use the app.

Several factors can influence the impact of mobile health apps on health outcomes, including user engagement, app design, and integration with existing healthcare systems. User engagement is critical for the effectiveness of mobile health apps, as users who are more engaged are more likely to adhere to the app's recommendations and achieve better health outcomes. App design is also important, as apps that are user-friendly, intuitive, and personalized are more likely to be effective. Additionally, the integration of mobile health apps with existing healthcare systems can enhance their impact by allowing for seamless communication between patients and healthcare providers and facilitating data sharing for better decision-making (Grundy, 2022, Qudah & Luetsch, 2019, Szinay, et. al., 2020).

In conclusion, mobile health apps have the potential to significantly improve health outcomes by providing users with access to health information, monitoring tools, and behavior change support. The evidence supporting the effectiveness of these apps is growing, with studies demonstrating their positive impact on a wide range of health conditions. Factors such as user engagement, app design, and integration with existing healthcare systems can influence the impact of mobile health apps on health outcomes, highlighting the need for continued research and innovation in this field.

5. Challenges and Limitations of Mobile Health Applications

Mobile health applications (apps) have the potential to revolutionize healthcare delivery by providing users with convenient access to health information, monitoring tools, and behavior change support. However, these apps also face a number of challenges and limitations that need to be addressed in order to maximize their effectiveness and impact. This review explores some of the key challenges and limitations of mobile health apps, including privacy and security concerns, reliability and accuracy of health information, and user engagement and adherence. One of the major

challenges facing mobile health apps is privacy and security concerns. Mobile health apps collect and store sensitive health information, such as medical history, medication lists, and biometric data Aljedaani, et. al., 2021, (Katarahweire, Bainomugisha & Mughal, 2020, Zhou, DeAlmeida & Parmanto, 019). This information is often transmitted over the internet, making it vulnerable to security breaches and unauthorized access.

To address these concerns, developers of mobile health apps need to implement robust security measures, such as encryption, secure authentication, and data anonymization. They also need to comply with relevant regulations, such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States, which sets standards for the protection of health information. Another challenge facing mobile health apps is the reliability and accuracy of the health information they provide. While many apps claim to provide evidence-based information and recommendations, there is often limited scientific evidence to support their claims. This can lead to misinformation and potentially harmful advice being provided to users.

To address this challenge, developers of mobile health apps should ensure that the information and recommendations they provide are based on sound scientific evidence. They should also make it clear to users the level of evidence supporting their recommendations and provide references to the sources of information used. A third challenge facing mobile health apps is user engagement and adherence. Many users download health apps with good intentions but fail to engage with them over the long term. This can be due to a variety of factors, including the complexity of the app, lack of motivation, and competing priorities.

To address this challenge, developers of mobile health apps should focus on designing apps that are user-friendly, intuitive, and engaging. They should also incorporate features that encourage regular use, such as reminders, rewards, and social networking features. Additionally, they should provide users with personalized feedback and recommendations based on their individual health goals and preferences.

In conclusion, mobile health apps have the potential to improve healthcare delivery and empower users to take control of their health. However, they also face a number of challenges and limitations, including privacy and security concerns, reliability and accuracy of health information, and user engagement and adherence. By addressing these challenges, developers of mobile health apps can maximize their effectiveness and impact, ultimately leading to better health outcomes for users.

5.1. Strategies for Enhancing the Role of Mobile Health Applications

Mobile health applications (apps) have the potential to revolutionize healthcare delivery by providing users with convenient access to health information, monitoring tools, and behavior change support. However, to maximize their effectiveness and impact, it is important to implement strategies to enhance their role. This review explores some key strategies for enhancing the role of mobile health apps, including improving user interface and usability, integrating apps into existing healthcare systems, and providing incentives for app usage and engagement.

One of the most important strategies for enhancing the role of mobile health apps is to improve their user interface and usability (Islam, et. al., 2020, Liew, et. al. 2019, Wei, et. al., 2020). Apps that are easy to navigate, visually appealing, and intuitive to use are more likely to be adopted and used regularly by users. This can be achieved through user-centered design principles, such as conducting user research, creating personas, and iterative testing with real users. Another key strategy for enhancing the role of mobile health apps is to integrate them into existing healthcare systems. This can be done by ensuring that apps comply with relevant regulations and standards, such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States. It can also involve integrating apps with electronic health records (EHRs) and other healthcare systems to enable seamless communication and data sharing between patients and healthcare providers.

Incentives can play a crucial role in encouraging users to download and engage with mobile health apps. This can include offering rewards, such as discounts on healthcare services or products, for reaching health goals or using the app regularly. Incentives can also be social in nature, such as allowing users to share their achievements with friends and family on social media (Amagai, et. al., 2022, Brower, et. al., 2020, Mahmood, et. al., 2019).

In conclusion, mobile health apps have the potential to revolutionize healthcare delivery by providing users with convenient access to health information, monitoring tools, and behavior change support. By implementing strategies to enhance their role, such as improving user interface and usability, integrating apps into existing healthcare systems, and providing incentives for app usage and engagement, developers and healthcare providers can maximize the effectiveness and impact of mobile health apps, ultimately leading to better health outcomes for users.

6. Case Studies

Mobile health applications (apps) have been instrumental in improving patient engagement and health outcomes in diabetes management (Batch, et. al., 2021, Cucciniello, et. al., 2021, Qudah & Luetsch, 2019). One such app is MySugr, which helps patients with diabetes track their blood glucose levels, medication intake, and physical activity. The app provides personalized feedback and insights based on the data entered by the user, empowering them to make informed decisions about their health. A study published in the Journal of Diabetes Science and Technology evaluated the impact of MySugr on glycemic control and self-management behaviors in patients with type 2 diabetes. The study found that patients who used the app experienced significant improvements in their HbA1c levels compared to those who did not use the app. Additionally, patients reported increased adherence to their medication regimens and dietary recommendations, leading to better overall health outcomes (Debong, Mayer & Kober, 2019, Kordonouri & Riddell, 2019, Maharaj, et. al., 2021).

Mobile health apps have also been effective in improving patient engagement and health outcomes in mental health care. The app Headspace, for example, offers guided meditation and mindfulness exercises to help users manage stress, anxiety, and depression. By providing accessible and personalized mental health support, Headspace has helped users improve their emotional well-being and quality of life. A study published in the Journal of Medical Internet Research examined the impact of Headspace on stress and anxiety levels in college students (Lipschitz, et. al., 2019, Ng, et. al., 2019, Zhang, et. al., 2019). The study found that students who used the app experienced significant reductions in stress and anxiety compared to those who did not use the app. Additionally, students reported improvements in their sleep quality and overall well-being, highlighting the positive impact of mobile health apps on mental health outcomes.

Mobile health apps have also shown promise in improving patient engagement and health outcomes in chronic disease management. The app Mango Health, for example, helps patients track their medication intake, set reminders, and receive rewards for adhering to their medication regimens. By gamifying the medication adherence process, Mango Health motivates patients to stay on track with their treatment plans. A study published in the Journal of Medical Internet Research studied the impact of Mango Health on medication adherence and health outcomes in patients with hypertension. The study found that patients who used the app were more likely to adhere to their medication regimens and achieve better blood pressure control compared to those who did not use the app. Additionally, patients reported increased motivation to manage their hypertension and improved overall health outcomes (Bardhan, Chen & Karahanna, 2020, Pham, et. al., 2019, Stewart, et. al., 2020).

Overall, these case studies demonstrate the significant impact of mobile health applications in improving patient engagement and health outcomes across various health conditions. By providing accessible, personalized, and engaging health support, mobile health apps empower patients to take control of their health and achieve better outcomes.

7. Future Directions and Opportunities

Mobile health applications (apps) have transformed healthcare delivery by providing users with convenient access to health information, monitoring tools, and behavior change support. As technology continues to evolve, there are several emerging trends and opportunities for mobile health apps to further improve patient engagement and health outcomes (Galetsi, Katsaliaki & Kumar, 2022, Pires, et. al., 2020, Wu, et. al., 2020). This review explores these future directions and opportunities, including emerging trends in mobile health app development, the potential for integrating artificial intelligence (AI) and machine learning (ML) into apps, and the implications for healthcare policy and practice.

One of the emerging trends in mobile health app development is the focus on personalized health solutions. Developers are increasingly incorporating features that allow users to track their health metrics, receive personalized recommendations, and connect with healthcare providers (Loncar-Turukalo, et. al., 2019, Rajak & Shaw, 2019). For example, apps that use wearable devices to track physical activity and provide personalized fitness plans are becoming increasingly popular. Another emerging trend is the use of gamification to enhance user engagement. Gamification involves incorporating game-like elements, such as rewards, challenges, and leaderboards, into apps to motivate users to engage with them regularly. This can be particularly effective for apps designed to promote healthy behaviors, such as exercise and diet tracking apps (Cheng, et. al., 2019, Saleem, Noori & Ozdamli, 2022, Tobon, Ruiz-Alba & García-Madariaga, 2020).

AI and ML have the potential to revolutionize mobile health apps by enabling them to provide more personalized and accurate health recommendations (Johnson, et. al., 2021, Khan & Alotaibi, 2020, Sarker, et. al., 2021). For example, AI algorithms can analyze user data, such as health metrics and lifestyle habits, to provide personalized health

recommendations and identify early signs of health problems. ML can also be used to improve the accuracy of health information provided by apps. For example, ML algorithms can analyze large datasets of medical literature to provide users with up-to-date and evidence-based health information.

The increasing popularity and effectiveness of mobile health apps have several implications for healthcare policy and practice. For example, policymakers may need to develop regulations to ensure the privacy and security of user data collected by health apps. They may also need to consider how to integrate mobile health apps into existing healthcare systems to improve communication between patients and healthcare providers. Healthcare providers may need to adapt their practices to accommodate the use of mobile health apps by patients. This may include providing training and support to patients on how to use health apps effectively and integrating app data into patient health records (Akbar, Coiera & Magrabi, 2020, Balapour, et. al., 2019, Wang & Qi, 2021).

In conclusion, mobile health apps have the potential to significantly improve patient engagement and health outcomes. By incorporating emerging trends in app development, such as personalized health solutions and gamification, and leveraging AI and ML technologies, developers can further enhance the effectiveness of mobile health apps. However, policymakers and healthcare providers must also address the challenges posed by mobile health apps, such as privacy and security concerns, to ensure that they can be effectively integrated into healthcare systems.

8. Conclusion

Mobile health applications (apps) have emerged as powerful tools for improving patient engagement and health outcomes. This critical review has highlighted the key findings on the role of mobile health apps in patient engagement and health outcomes, discussed recommendations for future research and practice, and provided closing remarks on the potential impact of mobile health apps on healthcare delivery. Mobile health apps play a crucial role in promoting patient engagement by providing users with access to health information, monitoring tools, and behavior change support. They empower patients to take control of their health and make informed decisions about their care. Studies have demonstrated that mobile health apps can lead to improved health outcomes, including better management of chronic conditions, increased medication adherence, and improved patient-provider communication.

To further enhance the role of mobile health apps in improving patient engagement and health outcomes, several recommendations can be made. First, there is a need for more research to evaluate the effectiveness of different types of mobile health apps in diverse patient populations and healthcare settings. Second, developers should focus on improving the user interface and usability of apps to enhance user engagement and adherence. Third, healthcare providers should integrate mobile health apps into their practice and provide support and guidance to patients on how to use them effectively.

Mobile health apps have the potential to significantly impact healthcare delivery by improving patient engagement and health outcomes. However, challenges such as privacy and security concerns, reliability and accuracy of health information, and user engagement and adherence need to be addressed to maximize their effectiveness. By addressing these challenges and implementing the recommendations outlined in this review, mobile health apps can become valuable tools for enhancing patient engagement and improving health outcomes across diverse populations.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

- [1] Akbar, S., Coiera, E., & Magrabi, F. (2020). Safety concerns with consumer-facing mobile health applications and their consequences: a scoping review. *Journal of the American Medical Informatics Association*, 27(2), 330-340.
- [2] Aljedaani, B., Ahmad, A., Zahedi, M., & Babar, M. A. (2021). End-users' knowledge and perception about security of mobile health apps: An empirical study. *arXiv preprint ArXiv:2101.10412 v3*.
- [3] Amagai, S., Pila, S., Kaat, A. J., Nowinski, C. J., & Gershon, R. C. (2022). Challenges in participant engagement and retention using mobile health apps: literature review. *Journal of medical Internet research*, 24(4), e35120.

- [4] Azizan, A., Ahmed, W., & Razak, A. H. A. (2023). Sensing health: a bibliometric analysis of wearable sensors in healthcare. *Health and Technology*, 1-20.
- [5] Balapour, A., Reychav, I., Sabherwal, R., & Azuri, J. (2019). Mobile technology identity and self-efficacy: Implications for the adoption of clinically supported mobile health apps. *International Journal of Information Management*, 49, 58-68.
- [6] Bardhan, I., Chen, H., & Karahanna, E. (2020). Connecting systems, data, and people: A multidisciplinary research roadmap for chronic disease management. *MIS Quarterly*, 44(1), 185-200.
- [7] Batch, B. C., Spratt, S. E., Blalock, D. V., Benditz, C., Weiss, A., Dolor, R. J., & Cho, A. H. (2021). General behavioral engagement and changes in clinical and cognitive outcomes of patients with type 2 diabetes using the Time2Focus mobile app for diabetes education: pilot evaluation. *Journal of Medical Internet Research*, 23(1), e17537.
- [8] Blaszka, M., & Rascon, N. A. (2023). Wearable fitness devices: An investigation into co-constructed meaning of use. *Communication & Sport*, 11(3), 570-591.
- [9] Bracken, M. L., & Waite, B. M. (2020). Self-efficacy and nutrition-related goal achievement of MyFitnessPal users. *Health Education & Behavior*, 47(5), 677-681.
- [10] Brower, J., LaBarge, M. C., White, L., & Mitchell, M. S. (2020). Examining responsiveness to an incentive-based Mobile health app: longitudinal observational study. *Journal of medical Internet research*, 22(8), e16797.
- [11] Cengiz, D., & Korkmaz, F. (2022). A Multidimensional Approach to Chronic Disease Management: The Patient Health Engagement Model. *Cyprus Journal of Medical Sciences*, 7(6).
- [12] Cheng, V. W. S., Davenport, T., Johnson, D., Vella, K., & Hickie, I. B. (2019). Gamification in apps and technologies for improving mental health and well-being: systematic review. *JMIR mental health*, 6(6), e13717.
- [13] Cucciniello, M., Petracca, F., Ciani, O., & Tarricone, R. (2021). Development features and study characteristics of mobile health apps in the management of chronic conditions: a systematic review of randomised trials. *NPJ Digital Medicine*, 4(1), 144.
- [14] Debon, R., Coleone, J. D., Bellei, E. A., & De Marchi, A. C. B. (2019). Mobile health applications for chronic diseases: A systematic review of features for lifestyle improvement. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 13(4), 2507-2512.
- [15] Debong, F., Mayer, H., & Kober, J. (2019). Real-world assessments of mySugr mobile health app. *Diabetes technology & therapeutics*, 21(S2), S2-35.
- [16] Galetsi, P., Katsaliaki, K., & Kumar, S. (2022). Assessing technology innovation of mobile health apps for medical care providers. *IEEE Transactions on Engineering Management*.
- [17] Galetsi, P., Katsaliaki, K., & Kumar, S. (2023). Exploring benefits and ethical challenges in the rise of mHealth (mobile healthcare) technology for the common good: An analysis of mobile applications for health specialists. *Technovation*, 121, 102598.
- [18] Ghose, A., Guo, X., Li, B., & Dang, Y. (2021). Empowering patients using smart mobile health platforms: Evidence from a randomized field experiment. *arXiv preprint arXiv:2102.05506*.
- [19] Gordon, M., Althoff, T., & Leskovec, J. (2019, May). Goal-setting and achievement in activity tracking apps: a case study of MyFitnessPal. In *The World Wide Web Conference* (pp. 571-582).
- [20] Graffigna, G., Barello, S., Riva, G., Corbo, M., Damiani, G., Iannone, P., ... & Ricciardi, W. (2020). Italian consensus statement on patient engagement in chronic care: process and outcomes. *International Journal of Environmental Research and Public Health*, 17(11), 4167.
- [21] Grundy, Q. (2022). A review of the quality and impact of mobile health apps. *Annual review of public health*, 43, 117-134.
- [22] Iribarren, S. J., Akande, T. O., Kamp, K. J., Barry, D., Kader, Y. G., & Suelzer, E. (2021). Effectiveness of mobile apps to promote health and manage disease: systematic review and meta-analysis of randomized controlled trials. *JMIR mHealth and uHealth*, 9(1), e21563.
- [23] Islam, M. N., Karim, M. M., Inan, T. T., & Islam, A. N. (2020). Investigating usability of mobile health applications in Bangladesh. *BMC medical informatics and decision making*, 20, 1-13.

- [24] Johnson, K. B., Wei, W. Q., Weeraratne, D., Frisse, M. E., Misulis, K., Rhee, K., ... & Snowdon, J. L. (2021). Precision medicine, AI, and the future of personalized health care. *Clinical and translational science*, 14(1), 86-93.
- [25] Katarahweire, M., Bainomugisha, E., & Mughal, K. A. (2020). Data classification for secure mobile health data collection systems. *Development Engineering*, 5, 100054.
- [26] Khan, Z. F., & Alotaibi, S. R. (2020). Applications of artificial intelligence and big data analytics in m-health: a healthcare system perspective. *Journal of healthcare engineering*, 2020, 1-15.
- [27] Kordonouri, O., & Riddell, M. C. (2019). Use of apps for physical activity in type 1 diabetes: current status and requirements for future development. *Therapeutic advances in endocrinology and metabolism*, 10, 2042018819839298.
- [28] Liew, M. S., Zhang, J., See, J., & Ong, Y. L. (2019). Usability challenges for health and wellness mobile apps: mixed-methods study among mHealth experts and consumers. *JMIR mHealth and uHealth*, 7(1), e12160.
- [29] Lipschitz, J., Miller, C. J., Hogan, T. P., Burdick, K. E., Lippin-Foster, R., Simon, S. R., & Burgess, J. (2019). Adoption of mobile apps for depression and anxiety: cross-sectional survey study on patient interest and barriers to engagement. *JMIR mental health*, 6(1), e11334.
- [30] Loncar-Turukalo, T., Zdravevski, E., da Silva, J. M., Chouvarda, I., & Trajkovik, V. (2019). Literature on wearable technology for connected health: scoping review of research trends, advances, and barriers. *Journal of medical Internet research*, 21(9), e14017.
- [31] Maharaj, A., Lim, D., Murphy, R., & Serlachius, A. (2021). Comparing two commercially available diabetes apps to explore challenges in user engagement: randomized controlled feasibility study. *JMIR formative research*, 5(6), e25151.
- [32] Mahmood, A., Kedia, S., Wyant, D. K., Ahn, S., & Bhuyan, S. S. (2019). Use of mobile health applications for health-promoting behavior among individuals with chronic medical conditions. *Digital health*, 5, 2055207619882181.
- [33] Ng, M. M., Firth, J., Minen, M., & Torous, J. (2019). User engagement in mental health apps: a review of measurement, reporting, and validity. *Psychiatric Services*, 70(7), 538-544.
- [34] Nguyen, A. D., Frensham, L. J., Baysari, M. T., Carland, J. E., & Day, R. O. (2019). Patients' use of mobile health applications: what general practitioners think. *Family practice*, 36(2), 214-218.
- [35] Nussbaum, R., Kelly, C., Quinby, E., Mac, A., Parmanto, B., & Dicianno, B. E. (2019). Systematic review of mobile health applications in rehabilitation. *Archives of physical medicine and rehabilitation*, 100(1), 115-127.
- [36] Patel, M. L., Hopkins, C. M., Brooks, T. L., & Bennett, G. G. (2019). Comparing self-monitoring strategies for weight loss in a smartphone app: randomized controlled trial. *JMIR mHealth and uHealth*, 7(2), e12209.
- [37] Pham, Q., Graham, G., Carrion, C., Morita, P. P., Seto, E., Stinson, J. N., & Cafazzo, J. A. (2019). A library of analytic indicators to evaluate effective engagement with consumer mHealth apps for chronic conditions: scoping review. *JMIR mHealth and uHealth*, 7(1), e11941.
- [38] Pires, I. M., Marques, G., Garcia, N. M., Flórez-Revuelta, F., Ponciano, V., & Oniani, S. (2020). A research on the classification and applicability of the mobile health applications. *Journal of personalized medicine*, 10(1), 11.
- [39] Qudah, B., & Luetsch, K. (2019). The influence of mobile health applications on patient-healthcare provider relationships: a systematic, narrative review. *Patient education and counseling*, 102(6), 1080-1089.
- [40] Rajak, M., & Shaw, K. (2019). Evaluation and selection of mobile health (mHealth) applications using AHP and fuzzy TOPSIS. *Technology in Society*, 59, 101186.
- [41] Saleem, A. N., Noori, N. M., & Ozdamli, F. (2022). Gamification applications in E-learning: A literature review. *Technology, Knowledge and Learning*, 27(1), 139-159.
- [42] Sarker, I. H., Hoque, M. M., Uddin, M. K., & Alsanoosy, T. (2021). Mobile data science and intelligent apps: concepts, ai-based modeling and research directions. *Mobile Networks and Applications*, 26, 285-303.
- [43] Schneider-Kamp, A., & Askegaard, S. (2020). Putting patients into the centre: Patient empowerment in everyday health practices. *Health*, 24(6), 625-645.
- [44] Song, T. (2021). Application of Mobile Health Services to Support Patient Self-Management of Chronic Conditions.

- [45] Stewart, M. T., Hogan, T. P., Nicklas, J., Robinson, S. A., Purington, C. M., Miller, C. J., ... & Shimada, S. L. (2020). The promise of patient portals for individuals living with chronic illness: qualitative study identifying pathways of patient engagement. *Journal of medical Internet research*, 22(7), e17744.
- [46] Szinay, D., Jones, A., Chadborn, T., Brown, J., & Naughton, F. (2020). Influences on the uptake of and engagement with health and well-being smartphone apps: systematic review. *Journal of medical Internet research*, 22(5), e17572.
- [47] Tanniru, M. (2019). Engagement leading to empowerment—Digital innovation strategies for patient care continuity. *Journal of Hospital Management and Health Policy*, 3.
- [48] Tobiano, G., Jerofke-Owen, T., & Marshall, A. P. (2021). Promoting patient engagement: a scoping review of actions that align with the interactive care model. *Scandinavian Journal of Caring Sciences*, 35(3), 722-741.
- [49] Tobon, S., Ruiz-Alba, J. L., & García-Madariaga, J. (2020). Gamification and online consumer decisions: Is the game over?. *Decision Support Systems*, 128, 113167.
- [50] Vaghefi, I., & Tulu, B. (2019). The continued use of mobile health apps: insights from a longitudinal study. *JMIR mHealth and uHealth*, 7(8), e12983.
- [51] Waalen, J. (2023). Mobile Health and Preventive Medicine. *Medical Clinics*, 107(6), 1097-1108.
- [52] Wang, C., & Qi, H. (2021, March). Influencing factors of acceptance and use behavior of mobile health application users: systematic review. In *Healthcare* (Vol. 9, No. 3, p. 357). MDPI.
- [53] Wei, Y., Zheng, P., Deng, H., Wang, X., Li, X., & Fu, H. (2020). Design features for improving mobile health intervention user engagement: systematic review and thematic analysis. *Journal of medical Internet research*, 22(12), e21687.
- [54] Wu, P., Zhang, R., Zhu, X., & Liu, M. (2022, January). Factors influencing continued usage behavior on mobile health applications. In *Healthcare* (Vol. 10, No. 2, p. 208). MDPI.
- [55] Zhang, R., Nicholas, J., Knapp, A. A., Graham, A. K., Gray, E., Kwasny, M. J., ... & Mohr, D. C. (2019). Clinically meaningful use of mental health apps and its effects on depression: mixed methods study. *Journal of medical Internet research*, 21(12), e15644.
- [56] Zhou, L., DeAlmeida, D., & Parmanto, B. (2019). Applying a user-centered approach to building a mobile personal health record app: development and usability study. *JMIR mHealth and uHealth*, 7(7), e13194.



Review

Exploring the Role of ChatGPT in Oncology: Providing Information and Support for Cancer Patients

Maurizio Cè ¹, Vittoria Chiарpenello ¹, Alessandra Bubba ¹, Paolo Florent Felisaz ², Giancarlo Oliva ², Giovanni Irmici ³ and Michaela Cellina ^{2,*}

¹ Postgraduation School in Radiodiagnostics, Università degli Studi di Milano, Via Festa del Perdono, 7, 20122 Milan, Italy; maurizio.ce@unimi.it (M.C.); vittoria.chiarpenello@unimi.it (V.C.)

² Radiology Department, Fatebenefratelli Hospital, ASST Fatebenefratelli Sacco, Piazza Principessa Clotilde 3, 20121 Milan, Italy

³ Radiology Department, IRCCS Istituto Nazionale dei Tumori, Via Giacomo Venezian, 1, 20133 Milan, Italy

* Correspondence: michaela.cellina@asst-fbf-sacco.it

Abstract: Introduction: Oncological patients face numerous challenges throughout their cancer journey while navigating complex medical information. The advent of AI-based conversational models like ChatGPT (San Francisco, OpenAI) represents an innovation in oncological patient management. Methods: We conducted a comprehensive review of the literature on the use of ChatGPT in providing tailored information and support to patients with various types of cancer, including head and neck, liver, prostate, breast, lung, pancreas, colon, and cervical cancer. Results and Discussion: Our findings indicate that, in most instances, ChatGPT responses were accurate, dependable, and aligned with the expertise of oncology professionals, especially for certain subtypes of cancers like head and neck and prostate cancers. Furthermore, the system demonstrated a remarkable ability to comprehend patients' emotional responses and offer proactive solutions and advice. Nevertheless, these models have also showed notable limitations and cannot serve as a substitute for the role of a physician under any circumstances. Conclusions: Conversational models like ChatGPT can significantly enhance the overall well-being and empowerment of oncological patients. Both patients and healthcare providers must become well-versed in the advantages and limitations of these emerging technologies.



Citation: Cè, M.; Chiарpenello, V.; Bubba, A.; Felisaz, P.F.; Oliva, G.; Irmici, G.; Cellina, M. Exploring the Role of ChatGPT in Oncology: Providing Information and Support for Cancer Patients. *BioMedInformatics* **2024**, *4*, 877–888. <https://doi.org/10.3390/biomedinformatics4020049>

Academic Editor: Carson K. Leung

Received: 15 January 2024

Revised: 19 March 2024

Accepted: 21 March 2024

Published: 25 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Oncological patients face numerous challenges throughout their cancer journey, ranging from emotional distress and treatment-related side effects, to navigating complex medical information.

Gone are the days when patients depended only on their doctors for medical advice. With a simple internet search, patients may educate themselves on symptoms, diseases, and treatment options, becoming more informed and proactive in decisions regarding their health [1].

The Internet has definitely transformed how patients navigate medical information, reshaping the dynamics of patient empowerment and communication with the doctor [2]. However, while access to medical knowledge can be useful under certain circumstances not all online sources are reliable, and patients may encounter incorrect or misleading information resulting in confusion or wrong self-diagnosis [3]. For this reason, misinformation or harmful information about cancer continues to be a significant concern in the online communication environment [4].

Furthermore, the ever-increasing health budget limits and heightened workloads among healthcare professionals have exacerbated the decline in doctor–patient relationships, adversely affecting healthcare accessibility and prognosis [5].

Providing cancer patients with additional tools for a better understanding of their diagnosis and treatment options and adequate emotional support is critical to ensure informed decision-making and a good outcome.

In this scenario, the advent of large language models (LLMs) like ChatGPT (San Francisco, OpenAI) and others may represent a cutting-edge innovation in oncological patient management to meet their individualized needs and concerns [6].

The use of artificial intelligence in healthcare is not new, having already demonstrated surprising results in the high-performance analysis of biomedical data through machine learning and deep learning models [7]. However, despite the great prospects, some issues related to reliability, privacy, and patient confidentiality should still be addressed when integrating these tools into healthcare routines [6,8–10]. This narrative review explores the potential advantages, limitations, and challenges associated with conversational models in supporting cancer patients. Our discussion includes aspects such as the accessibility of the models, the reliability of the information provided, as well as their role in patient empowerment and informed decision-making. We focus on the widely recognized large language model ChatGPT (developed by OpenAI, San Francisco) due to the availability of consistent literature on this topic [6].

Large Language Models

Large language models (LLMs) are sophisticated artificial intelligence systems designed to generate human-like text. They are trained on vast amounts of data and can understand and produce natural language across various tasks, such as translation, summarization, and conversation. Users provide a list of keywords or inquiries, and LLMs generate content about those topics. The user interface generally follows a conversational structure, which cycles between user questions or inputs and system responses or outputs. This design considers previous interactions to emulate human speech effectively [6].

In November 2022, OpenAI launched its GPT series of models (e.g., GPT-3.5, GPT-4, and GPT-5), which generate human-like text for usage in chatbot conversations using natural language processing (NLP) technology. Other notable LLMs are Google's PaLM and Gemini, Meta's LLaMA family of open-source models, Anthropic's Claude models, ChatBot BARD by Google, and Llama and Llama-2 by Meta.

LLMs are AI-driven, deep neural network-based models with a remarkable ability to achieve general-purpose human language generation and understanding [6–9]. LLMs acquire these skills by learning statistical relationships from text documents in computationally intensive self-supervised and semi-supervised training processes. Generative pre-trained transformer (GPT) language models are built on a transformer architecture, which enables them to process large amounts of text data while producing coherent text outputs by learning the relationships between input and output sequences [10].

The GPT language model has been trained on large datasets of text sourced from websites, books, and online publications. After receiving human feedback and corrections, ChatGPT was trained to respond in a way that allowed it to produce more logical and contextually relevant answers [11]; this procedure is known as reinforcement learning from human feedback or reinforcement learning from human preference (RLHF/RLHP). Users can type any prompt, and ChatGPT will answer based on the data stored in its database.

Previous research demonstrated that it could produce high-quality and coherent text outputs, react to user questions with unexpectedly intelligent-sounding messages, and perform exceptionally well in question-answer tasks [12]. In the medical area, GPT-4, created by further reinforcement learning from OpenAI's ChatGPT, recently surpassed the passing score on all steps of the US medical licensing exam [13].

2. Methods

An extensive literature search was performed on PubMed to find relevant publications on the current role and future potential of ChatGPT in cancer patients. For our search, we have the following string: “(cancer OR oncology OR oncological) AND (patients) AND

ChatGPT". Furthermore, we also carried out a careful examination of the references of the included articles to evaluate further studies worthy of mention.

Our results are presented through a narrative summary and organized as follows: potential benefits, applications in different types of cancer, limitations, and challenges.

3. Results and Discussion

The potential advantages and limitations of ChatGPT and similar LLMs are presented in Table 1.

Table 1. Advantages and limitations of ChatGPT and similar LLMs for patients and doctors.

Advantages	Limitations
Accessibility and inclusiveness (remote access to essential health information for disadvantaged communities, inability to provide harmful or offensive responses, and elimination of stigma surrounding sensitive topics)	Limited Internet access and low digital literacy. Some languages are not available.
Timely and accurate access to medical information (good for general purposes and basic medical knowledge)	Limited consideration of the patient's medical history; these models work only on inputs provided and cannot ask questions (moderate accuracy for complex cases).
Patient-friendly explanations of medical terms, treatment options, and potential side effects to improve patient understanding and informed decision-making	Oversimplification. Patients may consider this tool as a substitute for medical consultation, replacing the doctor–patient relationship.
Emotional support and patient empowerment	Underestimation of disease severity.
Ethical and legal support	It may not comply with local medical regulations, standards, and the patient's cultural background.
Handling routine repetitive tasks like writing medical reports	Reduced supervision of clinicians routinely relies on these tools.
Retrieval of medical literature	It does not provide references (or if it does, they are not always true). It is limited to information available until the knowledge cutoff date and does not have real-time updates. It may include biased or outdated information. Limited consideration of the specific clinical setting.

3.1. Potential Advantages of ChatGPT and Other LLMs

3.1.1. Accessibility and Inclusivity

LLMs represent a potential game changer for individuals with limited access to health-care resources, particularly in low-income countries, as, despite widespread access to medical information contributing to increasing the average level of health literacy and well-being expectations, health services are still very unequal and insufficient in many parts of the world [14,15].

As of the time this article was written, the basic version of ChatGPT was free of charge for the public. Given that financial difficulties have been linked to poor health outcomes [16], large language models can contribute to limiting the effects of socioeconomic inequalities in cancer treatment by giving everyone fast access to reliable medical information regardless of their location or socioeconomic background [17–20].

ChatGPT can support underprivileged communities in many ways: first, ChatGPT communicates in multiple languages, breaking down language barriers that often hinder access to healthcare information. It can remotely deliver essential health information in areas with limited access to healthcare facilities and guide essential self-care practices, including managing chronic conditions and first aid measures. In developing countries, it can help individuals understand healthcare processes, such as insurance enrollment, appointment scheduling, and medication management, improving overall access to care. Overall, ChatGPT has the potential to democratize access to healthcare information and support, ultimately improving health outcomes for underprivileged communities [17–20].

The inability of ChatGPT to generate any offensive or harmful responses is one of the security measures put in place by the developers to prevent misuse. ChatGPT can, therefore, offers a non-judgmental platform to seek information on sensitive topics as well, such as sexual health, mental health, and substance abuse, reducing stigma and cultural barriers that often deter people from seeking help.

3.1.2. Information Provision and Informed Decision-Making

ChatGPT has been trained on a large amount of data, including medical literature [7]. Even if, as discussed further, the input (training) data constitute a limiting factor for deep learning models' accuracy and trustworthiness, these tools represent a valuable supplement for medical information retrieval and clinical decision-making, both for patients and healthcare practitioners. ChatGPT's conversational style results in more comprehensible responses than primary official sources like guidelines and scientific articles, especially for individuals without expertise in the medical field. Additionally, it streamlines the information search process by presenting only relevant content tailored to the user's query, thus enhancing efficiency and saving time.

LLMs are already active in different areas of clinical practice and can generate differential diagnosis lists for typical clinical scenarios with good diagnostic accuracy [12]. In oncology, by integrating this knowledge into coherent responses, ChatGPT can answer questions related to different types of cancer, including treatment options, potential side effects, and beneficial lifestyle modifications. These models can support patients asking for information about additional examinations, diagnosis, treatment plans, and prognosis, enabling them to make more informed decisions. In breast cancer imaging, for example, it performs reasonably well in recommending the next imaging steps for patients requiring a breast cancer screening and assessment of breast pain [18]. The adequacy of ChatGPT and other LLMs as a guide for patients, who are non-experts in medicine, in navigating the correct diagnostic path, remains a contentious issue in ongoing discussions.

Additionally, ChatGPT can assist in clarifying medical terminology and lexicon, ensuring that patients better comprehend the information provided in medical documents and radiological reports [19].

3.1.3. Emotional Support and Patient Empowerment

A cancer diagnosis can often lead to emotional distress, anxiety, and depression for patients and their caregivers [20,21]. Furthermore, because of time constraints, clinician-patient communication may frequently be neglected, with dramatic consequences for the clinical history and life of cancer patients.

There are good reasons to think that ChatGPT could help bridge this gap.

Research into ChatGPT's ability to provide responses attuned to human emotions such as kindness and empathy has produced impressive results [22]. It may give the impression that generative AI can demonstrate an understanding of human emotions, generating responses and assistance suitable for those who use it. In a recent study by Elyoseph et al. [23], ChatGPT outperformed humans in assessing emotional awareness. It demonstrated the ability to improve intrapersonal and interpersonal understanding, increasing patients' awareness of their own and their family members' emotions. This may provide patients with comfort and help them feel less "alone" [24].

Using natural language processing capabilities, ChatGPT can engage in compassionate conversations, acknowledging patients' emotions and providing emotional support [25]. It can also suggest coping strategies, stress management techniques, and even provide referrals to mental health professionals, when necessary.

It helps to build a framework for each individual question presented by patients and caregivers, thus increasing provider efficiency and allowing patients to become more aware about their care. As a result, by providing patients with an additional source of information, this paradigm has the potential to boost patient participation and compliance, promoting patient-centered treatment and effective shared decision-making [11].

3.1.4. Supportive Care

Beyond medical treatment, oncological patients often require support in other aspects of their lives, such as managing relationships, and making lifestyle changes to preserve their health.

By offering suggestions for healthy lifestyle modifications, including exercise routines and dietary recommendations, ChatGPT can empower patients to take an active role in their overall well-being [17].

This model has showcased significant potential in aiding home care for orthopaedic patients, suggesting that this tool can play a pivotal role in improving public health policy by providing consistent and trustworthy guidance, especially in settings where access to health services is limited [26].

3.1.5. For Healthcare Practitioners and Medical Students

Tools like ChatGPT could be helpful not only for patients but also for healthcare practitioners [27]. Conversational models can generate user-friendly explanations of medical jargon, treatment alternatives, and potential adverse effects, thereby improving patient literacy and decision-making. Of course, ChatGPT is better suited to text activities like generating summaries, treatment plans, and follow-up recommendations, which doctors may subsequently check. Furthermore, it can facilitate contact with patients from varied linguistic backgrounds by offering real-time language translation services during consultations.

Another potentially beneficial use of tools like ChatGPT is training medical students and residents by simulating patient scenarios, answering medical queries, and providing learning resources [28,29].

Advanced AI-based models could save time for oncologists by handling routine administrative tasks like scheduling appointments, sending reminders, and managing documentation. It can help with patient case documentation by creating summaries of consultations, treatment plans, and follow-up recommendations and streamlining the process of keeping complete and accurate patient records. ChatGPT can generate clinic letters with good overall correctness and humanness ratings, with a reading level roughly similar to current real-world human-generated letters, and it has been effectively used to create patient clinic letters [30].

However, much caution should likely be exercised when considering specific tasks such as information retrieval about the latest research, treatment guidelines, clinical trials related to particular types of cancer, drug interactions, side effects, and dosage information for various cancer medications. Recent studies have shown a lack of consistency when dealing with providing a threshold for decision-making or distinguishing which guidelines to follow in a specific setting [11].

3.2. Appraisal of Literature on Different Types of Cancer

So far, the literature evaluating its use in clinical practice is still limited [8] and only a few studies have evaluated the potential of ChatGPT in education and advice on the clinical path of oncology [11].

3.2.1. Head and Neck

Kuṣcu et al. explored the accuracy and reliability of ChatGPT's responses to questions related to head and neck cancer [31]. A dataset of questions was selected from commonly asked queries from reputable institutions and societies, including the American Head & Neck Society (AHNS), the National Cancer Institute, and the Medline Plus Medical Encyclopedia. These questions underwent an extensive screening process by three authors to determine their suitability for inclusion in the study, focusing primarily on patient-oriented questions to evaluate the effectiveness of the AI model in providing useful information for patients. The study revealed that the majority of ChatGPT responses were accurate, with 86.4% receiving a "complete/correct" rating on the rating scale. Significantly, none of the responses were rated "completely inaccurate/irrelevant". Furthermore, the

model showed high reproducibility across all topics and performed consistently without significant differences between them.

The authors also underlined a substantial limitation of ChatGPT: the version knowledge cutoff was only extended until September 2021, potentially impacting response precision due to the exclusion of data from the previous two years. Furthermore, the reliability of ChatGPT is determined by the quality of its training data, and the model's secret sources raise questions about whether the training was based on the most reputable and accurate medical literature. Furthermore, the latest version of ChatGPT, which demonstrated better performance than the publicly available version, is exclusively accessible through paid membership, potentially restricting public access to more accurate knowledge [31].

A critical opinion regarding the current potential of ChatGPT in answering patient questions comes from the study by Wei et al., who compared the performance of ChatGPT and the Google browser in addressing common questions related to head and neck cancers [32]. A collection of 49 questions about head and neck cancers was chosen from a series of "People Also Ask" (PAA) question prompts using SearchResponse.io. The study found that, on average, Google sources outperformed ChatGPT responses. Both sources were assessed to be of similar readability difficulty, most likely at the college level. While ChatGPT responses were comparable in complexity to those from Google, they were rated as lower quality due to a drop in reliability and accuracy when answering questions.

According to Wei's assessment, particularly for questions about head and neck cancer, Google sources emerged as the primary option for patient educational resources [32].

3.2.2. Prostate Cancer

Zhu et al. developed a questionnaire aligning with patient education guidelines and their clinical expertise, covering screening, prevention, treatment options, and postoperative complications related to prostate cancer [17].

The questions covered a spectrum of knowledge from the basics to advanced knowledge about prostate cancer. Their investigation involved five Large Language Models, including ChatGPT (Free and Plus versions), YouChat, NeevaAI, Perplexity, and Chatsonic. Assessments revealed that LLMs excelled in addressing most questions. For instance, they effectively clarified the significance of different PSA levels and emphasized that PSA alone is not a conclusive diagnostic test and that further examinations are recommended. LLMs also demonstrated effectiveness in detailed comparisons of treatment options, presenting pros and cons, and offering informative references to aid patients in making well-informed decisions. Most importantly, in most cases, it was consistently emphasized to consult a doctor.

The accuracy of responses from most LLMs exceeded 90%, with exceptions noted for NeevaAI and Chatsonic. Basic information questions with definite answers generally achieved high accuracy, but accuracy dipped for queries tied to specific scenarios or requiring summarization and analysis. ChatGPT exhibited the highest accuracy rate among the LLMs assessed, with the free version slightly outperforming the paid version.

Zhu et al. raised a question in their study regarding whether online LLMs would surpass ChatGPT. Notably, AI models relying on search engines like NeevaAI often presented literature content without effective summarization and explanation, resulting in poor readability. This observation suggested that model training might be more crucial than real-time Internet connectivity [17].

3.2.3. Hepatocarcinoma

Individuals with cirrhosis and hepatocellular carcinoma (HCC), as well as their caregivers, often have unmet needs and insufficient knowledge regarding the management and prevention of complications associated with the disease. It should not be disregarded that a portion of these patients have a troublesome history behind them and lack a sufficient socioeconomic support network. Previous research has demonstrated inadequate

health literacy among cirrhosis and HCC patients and the favorable impacts of focused education [33].

An interesting experience comes from the work of Yeo et al., who evaluated ChatGPT's performance in answering the most frequently asked questions regarding the management and care of patients with cirrhosis and HCC. Conversational model responses were independently scored by two transplant hepatologists and a third reviewer [11].

The study by Yeo et al. found that ChatGPT provided comprehensive or correct but inadequate answers about cirrhosis in approximately three-quarters of the responses analyzed, with better results in categories such as "basic knowledge", "treatment", "lifestyle", and "other". No answer related to cirrhosis was classified as completely incorrect. Regarding HCC, the model excels in providing detailed information on the knowledge base and potential side effects of various HCC treatments, as well as scientific evidence for lifestyle investigations. However, there were areas where the model did not respond correctly or provided outdated information, especially in diagnosis, where most information was classified as a mix of correct and incorrect or outdated data. For example, while ChatGPT correctly emphasized using abdominal ultrasound as a primary screening tool, it neglected to mention MRI and computed tomography scans for HCC surveillance in patients with ascites. However, ChatGPT accurately identified cirrhosis as an indication for HCC surveillance [11].

Overall, the results were deemed satisfactory, even though only 47.3% of cirrhosis cases and 41.1% of HCC cases were classified as comprehensive, and the system had significant shortcomings in delivering answers about oncological diagnosis. Furthermore, the system could not establish the choice limitations and treatment length, most likely due to a lack of ability to deny clinical information regarding local procedures and recommendations. This confirms the potential significance of ChatGPT and related models in providing universal access to basic medical knowledge, while simultaneously emphasizing the importance of medical consultation during the most essential stages of the diagnostic process.

Yeo et al. also evaluated ChatGPT's responses to questions about coping with psychological stress following an HCC diagnosis. The model acknowledged the patient's probable emotional response to the diagnosis and provided clear and actionable starting points for individuals newly diagnosed with HCC. It offered motivational responses, encouraging proactive steps in managing the diagnosis and treatment strategies [11].

3.2.4. Breast Cancer

Over the last two decades, there has been an increase in scientific research and public interest in the two most serious problems linked with breast implants. Significant progress has been made in understanding the rare T-cell lymphoma associated with textured implants.

Liu et al. investigated the suitability of ChatGPT for educating patients on breast implant-associated anaplastic large cell lymphoma (BIA-ALCL) and breast implant illness (BII). They compared the quality of responses and references offered by ChatGPT to the Google Bard service. The data demonstrated that ChatGPT outperformed Google in providing high-quality responses to frequently asked queries about BIA-ALCL and BII [34].

3.2.5. Lung Cancer

Rahsepar et al. studied the accuracy of responses provided by ChatGPT-3.5, Google Bard, Bing, and Google search engines to non-expert questions about lung cancer prevention, screening, and vocabulary in radiology reports [35]. Out of 120 questions, ChatGPT-3.5 answered 70.8% correctly and 17.5% incorrectly. Google Bard did not respond to 23 queries, and of the 97 questions it did, 62 were correct, 11 had some errors, and 24 were incorrect. Out of 120 questions, Bing gave 61.7% correct, 10.8% mostly correct, and 27.5% incorrect answers. The Google search engine answered 120 questions with 55% correct, 22.5% mostly correct, and 22.5% incorrect.

The authors concluded that ChatGPT-3.5 was more likely to give correct or partially correct responses than Google Bard.

3.2.6. Colon Cancer

Regarding colon cancer, ChatGPT was twice asked 38 questions on prevention, diagnosis, and management, and three experts rated the appropriateness. Twenty-seven answers of ChatGPT were rated as “appropriate” by the three experts; overall, at least two of three experts rated the answers appropriate for 86.8% [36]. Moreover, the ChatGPT responses were extensively concordant with those of the American Society of Colon and Rectal Surgeons.

3.2.7. Pancreatic Cancer

Another study investigated the responses provided by ChatGPT to 30 questions about pancreatic cancer and pre-surgical, surgical, and post-surgical phases [37]. The response quality was then assessed by 20 surgical oncology experts and rated as ‘poor’, ‘fair’, ‘good’, ‘very good’, and ‘excellent’. The most assigned quality grade was ‘very good’ or ‘excellent’ for most responses ($n = 24/30, 80\%$); in total, 60% of the experts thought that ChatGPT was a reliable information source, and only 10% thought that the answers provided by ChatGPT could not be compared to those of skilled surgeons. Additionally, 90% of experts believed that ChatGPT will become the go-to source for online patient information, either completely replacing traditional search engines or at least co-existing with them.

3.2.8. Cervical Cancer

In a study by Hermann et al., when ChatGPT was challenged with questions concerning cervical cancer prevention, management, survivorship, and quality of life, its answers were rated as correct and comprehensive only in 34/64 (53.1%) questions, with the worst performance in the treatment category [38].

3.2.9. Radiotherapy

Although the authors did not use ChatGPT, the study by Chow et al. provides an instructive example of the efficiency of comparable conversational models. Their research focused on developing an AI-driven instructional chatbot for interactive learning in radiotherapy, using the IBM Watson Assistant platform [39].

The major purpose of the chatbot was to make it easier to communicate radiation knowledge to people of varied comprehension levels. The chatbot was created to be user-friendly and deliver simple explanations in response to user questions regarding radiation. According to their response, most physicians rated the RT Bot’s material positively, with 95% of users believing the information to be sufficiently complete.

4. Limitations and Perspectives

Healthcare professionals must be aware of the limitations of LLMs to ensure responsible and safe use.

Although ChatGPT is free and can benefit underprivileged communities who have difficulty accessing healthcare institutions, it is important to address constraints that remain in many parts of the world, such as limited Internet connection and low digital literacy.

Conversational models can be essential tools for physicians in providing general information and context, but they should not be relied on for medical advice. ChatGPT does not offer references (or if it does, they are not necessarily correct) [40]. Furthermore, it is limited to information available until the knowledge cutoff date. It does not have real-time updates, so it might not be aware of the latest medical breakthroughs, treatments, guidelines, or changes to regulations and laws. Since the model is trained on a diverse range of Internet texts, which may include biased or outdated information, this could lead to biased responses or recommendations that do not consider the most current and evidence-based medical practices. Different sources could reach different conclusions. This overlooks the current limitations in data accuracy, the evolving nature of medical knowledge, and the need for expert oversight.

From the patient's perspective, one of the potentially most harmful outcomes of the inappropriate use of ChatGPT is its ability to provide confidently stated yet incorrect answers [41], and it may be susceptible to what is termed "hallucinations", wherein information is fabricated rather than grounded in facts [42]. The average user often finds it more accessible to discern reliable sources, such as those affiliated with reputable health-care institutions or scientific organizations. Conversely, identifying erroneous information presented by ChatGPT can pose greater challenges due to its formal and plausible language delivery, coupled with the inability to trace its source. Future research could investigate how models like ChatGPT may inadvertently deceive not only individuals lacking medical training but also doctors who are not experts in the field compared to experts in the field.

General-purpose LLMs might not guarantee the accuracy and precision required for medical inquiries, which could lead to incorrect advice or information [43]. ChatGPT does not have access to personal health information about individuals. LLMs are not able to consider an individual's complete medical history, conduct physical examinations, or order diagnostic tests, which are essential aspects of providing advice for accurate and personalized medicine [9,44]. Any attempt to provide personalized medical advice would, therefore, be speculative and lead to inaccurate or potentially harmful recommendations.

Even while the quick availability of information helps to reduce anxiety, using LLM conversations without expert evaluation increases the danger of inaccuracy. For example, underestimating a patient's condition could negatively influence patient care, as erroneous results reporting or treatment guideline interpretation can affect patients' morbidity and mortality. Patients may develop a sense of comfort and trust with ChatGPT over time, contributing to enhanced emotional well-being. However, this sensation of comfort should not lead to an underestimation of the clinical state, causing the patient to make poor decisions. There is a real risk of oversimplifying complex medical situations, leading patients to believe the tool is a substitute for competent medical advice. Such a perception could undermine the crucial doctor–patient relationship founded on trust, expertise, and personalized care.

Therefore, while ChatGPT can support patient education, healthcare providers need to guide patients in using this tool as a complement to, rather than a substitute for, medical consultation. LLMs should be used carefully under the supervision of a qualified professional, oncologist, and psycho-oncologist to prevent the patient from forming incorrect beliefs about their illness.

Finally, providing medical advice involves legal and ethical considerations, and relying on a language model like ChatGPT may not comply with medical regulations, standards, or the patient's cultural background.

In conclusion, AI in healthcare must be strictly regulated and overseen to reduce these risks [6,30]. Further research is needed to compare the performance of different AI systems and evaluate the usefulness of AI-generated responses for cancer patients in real-world clinical settings. Seeking advice from experienced healthcare professionals who can assess individual clinical histories, conduct physical examinations, and interpret diagnostic testing is critical for providing accurate and safe medical care. Furthermore, it is vital to determine the quality and composition style of input delivered to chatbots across different settings, languages, and resource capacities. Implementing such a significant technological advancement necessitates caution and proactive risk management to ensure patient safety and quality of care.

5. Conclusions

The emergence of AI-driven conversational technology, exemplified by ChatGPT, has created new opportunities to support cancer patients throughout their journey. LLMs can significantly improve patients' well-being and empowerment by offering accurate information, guidance in treatment decisions, and emotional support. Evidence shows how these models can satisfactorily answer many questions about the symptoms, pathophysiology, treatment options, and prognosis of various types of cancer. However, these

models have limitations, the main concern being their potential to produce inaccurate or unreliable information plausibly, especially when dealing with complex medical conditions or nuanced treatment options. Additionally, ChatGPT may not interpret the context accurately or understand the subtle nuances of patient questions, leading to responses that are not fully applicable or helpful. Recognizing its limitations, integrating ChatGPT into the healthcare ecosystem promises to provide personalized, accessible, and empathetic support to cancer patients.

Author Contributions: Conceptualization, M.C. (Michaela Cellina) and M.C. (Maurizio Cè); methodology, V.C., A.B.; investigation, M.C. (Maurizio Cè), M.C. (Michaela Cellina) and G.O.; resources, V.C., A.B., P.F.F. and G.I.; writing—original draft preparation, M.C. (Maurizio Cè), M.C. (Michaela Cellina), V.C., A.B., P.F.F. and G.I.; writing—review and editing, G.O., M.C. (Maurizio Cè) and G.I.; supervision, P.F.F.; project administration, M.C. (Maurizio Cè) and M.C (Michaela Cellina). All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bujnowska-Fedak, M.M.; Waligóra, J.; Mastalerz-Migas, A. The Internet as a Source of Health Information and Services. In *Advancements and Innovations in Health Sciences*; Pokorski, M., Ed.; Advances in Experimental Medicine and Biology; Springer International Publishing: Cham, Switzerland, 2019; Volume 1211, pp. 1–16. [[CrossRef](#)]
2. Johnson, S.B.; King, A.J.; Warner, E.L.; Aneja, S.; Kann, B.H.; Bylund, C.L. Using ChatGPT to Evaluate Cancer Myths and Misconceptions: Artificial Intelligence and Cancer Information. *JNCI Cancer Spectr.* **2023**, *7*, pkad015. [[CrossRef](#)]
3. Yeung, A.W.K.; Tosevska, A.; Klager, E.; Eibenstein, F.; Tsagkaris, C.; Parvanov, E.D.; Nawaz, F.A.; Völk-Kernstock, S.; Schaden, E.; Kletecka-Pulker, M.; et al. Medical and Health-Related Misinformation on Social Media: Bibliometric Study of the Scientific Literature. *J. Med. Internet Res.* **2022**, *24*, e28152. [[CrossRef](#)] [[PubMed](#)]
4. Cancer Misinformation and Harmful Information on Facebook and Other Social Media: A Brief Report. Available online: <https://pubmed.ncbi.nlm.nih.gov/34291289/> (accessed on 21 February 2024).
5. Schäfer, W.L.A.; Van Den Berg, M.J.; Groenewegen, P.P. The Association between the Workload of General Practitioners and Patient Experiences with Care: Results of a Cross-Sectional Study in 33 Countries. *Hum. Resour. Health* **2020**, *18*, 76. [[CrossRef](#)] [[PubMed](#)]
6. Dave, T.; Athaluri, S.A.; Singh, S. ChatGPT in Medicine: An Overview of Its Applications, Advantages, Limitations, Future Prospects, and Ethical Considerations. *Front. Artif. Intell.* **2023**, *6*, 1169595. [[CrossRef](#)] [[PubMed](#)]
7. Joshi, G.; Jain, A.; Araveeti, S.R.; Adhikari, S.; Garg, H.; Bhandari, M. FDA-Approved Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices: An Updated Landscape. *Electronics* **2024**, *13*, 498. [[CrossRef](#)]
8. Li, Y.; Gao, W.; Luan, Z.; Zhou, Z.; Li, J. The Impact of Chat Generative Pre-Trained Transformer (ChatGPT) on Oncology: Application, Expectations, and Future Prospects. *Cureus* **2023**, *15*, e48670. [[CrossRef](#)] [[PubMed](#)]
9. Wang, C.; Liu, S.; Yang, H.; Guo, J.; Wu, Y.; Liu, J. Ethical Considerations of Using ChatGPT in Health Care. *J. Med. Internet Res.* **2023**, *25*, e48009. [[CrossRef](#)] [[PubMed](#)]
10. Chatterjee, J.; Dethlefs, N. This New Conversational AI Model Can Be Your Friend, Philosopher, and Guide . . . and Even Your Worst Enemy. *Patterns* **2023**, *4*, 100676. [[CrossRef](#)] [[PubMed](#)]
11. Hirosawa, T.; Harada, Y.; Yokose, M.; Sakamoto, T.; Kawamura, R.; Shimizu, T. Diagnostic Accuracy of Differential-Diagnosis Lists Generated by Generative Pretrained Transformer 3 Chatbot for Clinical Vignettes with Common Chief Complaints: A Pilot Study. *Int. J. Environ. Res. Public Health* **2023**, *20*, 3378. [[CrossRef](#)] [[PubMed](#)]
12. Kung, T.H.; Cheatham, M.; Medenilla, A.; Sillos, C.; De Leon, L.; Elepaño, C.; Madriaga, M.; Aggabao, R.; Diaz-Candido, G.; Maningo, J.; et al. Performance of ChatGPT on USMLE: Potential for AI-Assisted Medical Education Using Large Language Models. *PLoS Digit. Health* **2023**, *2*, e0000198. [[CrossRef](#)]
13. DeWalt, D.A.; Berkman, N.D.; Sheridan, S.; Lohr, K.N.; Pignone, M.P. Literacy and Health Outcomes: A Systematic Review of the Literature. *J. Gen. Intern. Med.* **2004**, *19*, 1228–1239. [[CrossRef](#)] [[PubMed](#)]
14. Budhathoki, S.S.; Pokharel, P.K.; Good, S.; Limbu, S.; Bhattachan, M.; Osborne, R.H. The Potential of Health Literacy to Address the Health Related UN Sustainable Development Goal 3 (SDG3) in Nepal: A Rapid Review. *BMC Health Serv. Res.* **2017**, *17*, 237. [[CrossRef](#)] [[PubMed](#)]
15. Perrone, F.; Jommi, C.; Di Maio, M.; Gimigliano, A.; Gridelli, C.; Pignata, S.; Ciardiello, F.; Nuzzo, F.; De Matteis, A.; Del Mastro, L.; et al. The Association of Financial Difficulties with Clinical Outcomes in Cancer Patients: Secondary Analysis of 16 Academic Prospective Clinical Trials Conducted in Italy. *Ann. Oncol.* **2016**, *27*, 2224–2229. [[CrossRef](#)]

16. Zhu, L.; Mou, W.; Chen, R. Can the ChatGPT and Other Large Language Models with Internet-Connected Database Solve the Questions and Concerns of Patient with Prostate Cancer and Help Democratize Medical Knowledge? *J. Transl. Med.* **2023**, *21*, 269. [CrossRef] [PubMed]
17. Rao, A.; Kim, J.; Kamineni, M.; Pang, M.; Lie, W.; Succi, M.D. Evaluating ChatGPT as an Adjunct for Radiologic Decision-Making. *MedRxiv* **2023**. preprint. [CrossRef]
18. Campbell, D.J.; Estephan, L.E. ChatGPT for Patient Education: An Evolving investigation. *J. Clin. Sleep Med.* **2023**, *19*, 2135–2136. [CrossRef] [PubMed]
19. Ryan, H.; Schofield, P.; Cockburn, J.; Butow, P.; Tattersall, M.; Turner, J.; Girgis, A.; Bandaranayake, D.; Bowman, D. How to Recognize and Manage Psychological Distress in Cancer Patients. *Eur. J. Cancer Care* **2005**, *14*, 7–15. [CrossRef] [PubMed]
20. Dekker, J.; Graves, K.D.; Badger, T.A.; Diefenbach, M.A. Management of Distress in Patients with Cancer—Are We Doing the Right Thing? *Ann. Behav. Med.* **2020**, *54*, 978–984. [CrossRef] [PubMed]
21. Gordijn, B.; Have, H.T. ChatGPT: Evolution or Revolution? *Med. Health Care Philos.* **2023**, *26*, 1–2. [CrossRef] [PubMed]
22. Elyoseph, Z.; Hadar-Shoval, D.; Asraf, K.; Lvovsky, M. ChatGPT Outperforms Humans in Emotional Awareness Evaluations. *Front. Psychol.* **2023**, *14*, 1199058. [CrossRef] [PubMed]
23. Dolan, N.C.; Ferreira, M.R.; Davis, T.C.; Fitzgibbon, M.L.; Rademaker, A.; Liu, D.; Schmitt, B.P.; Gorby, N.; Wolf, M.; Bennett, C.L. Colorectal Cancer Screening Knowledge, Attitudes, and Beliefs Among Veterans: Does Literacy Make a Difference? *J. Clin. Oncol.* **2004**, *22*, 2617–2622. [CrossRef] [PubMed]
24. Zheng, Y.; Wu, Y.; Feng, B.; Wang, L.; Kang, K.; Zhao, A. Enhancing Diabetes Self-Management and Education: A Critical Analysis of ChatGPT’s Role. *Ann. Biomed. Eng.* **2023**, *52*, 741–744. [CrossRef] [PubMed]
25. Yeo, Y.H.; Samaan, J.S.; Ng, W.H.; Ting, P.-S.; Trivedi, H.; Vipani, A.; Ayoub, W.; Yang, J.D.; Liran, O.; Spiegel, B.; et al. Assessing the Performance of ChatGPT in Answering Questions Regarding Cirrhosis and Hepatocellular Carcinoma. *Clin. Mol. Hepatol.* **2023**, *29*, 721–732. [CrossRef] [PubMed]
26. Yapar, D.; Demir Avci, Y.; Tokur Sonuvar, E.; Faruk Eğerci, Ö.; Yapar, A. ChatGPT’s Potential to Support Home Care for Patients in the Early Period after Orthopedic Interventions and Enhance Public Health. *Jt. Dis. Relat. Surg.* **2024**, *35*, 169–176. [CrossRef] [PubMed]
27. Borkowski, A.A. Applications of ChatGPT and Large Language Models in Medicine and Health Care: Benefits and Pitfalls. *Fed. Pract.* **2023**, *40*, 170. [CrossRef]
28. Tsang, R. Practical Applications of ChatGPT in Undergraduate Medical Education. *J. Med. Educ. Curric. Dev.* **2023**, *10*, 238212052311784. [CrossRef] [PubMed]
29. Khan, R.A.; Jawaad, M.; Khan, A.R.; Sajjad, M. ChatGPT—Reshaping Medical Education and Clinical Management. *Pak. J. Med. Sci.* **2023**, *39*. [CrossRef] [PubMed]
30. Kitamura, F.C. ChatGPT Is Shaping the Future of Medical Writing But Still Requires Human Judgment. *Radiology* **2023**, *307*, e230171. [CrossRef] [PubMed]
31. Kuşcu, O.; Pamuk, A.E.; Sütay Süslü, N.; Hosal, S. Is ChatGPT Accurate and Reliable in Answering Questions Regarding Head and Neck Cancer? *Front. Oncol.* **2023**, *13*, 1256459. [CrossRef]
32. Wei, K.; Fritz, C.; Rajasekaran, K. Answering Head and Neck Cancer Questions: An Assessment of ChatGPT Responses. *Am. J. Otolaryngol.* **2024**, *45*, 104085. [CrossRef]
33. Shaw, J.; Patidar, K.R.; Reuter, B.; Hajezifar, N.; Dharel, N.; Wade, J.B.; Bajaj, J.S. Focused Education Increases Hepatocellular Cancer Screening in Patients with Cirrhosis Regardless of Functional Health Literacy. *Dig. Dis. Sci.* **2021**, *66*, 2603–2609. [CrossRef] [PubMed]
34. Liu, H.Y.; Alessandri Bonetti, M.; De Lorenzi, F.; Gimbel, M.L.; Nguyen, V.T.; Egro, F.M. Consulting the Digital Doctor: Google Versus ChatGPT as Sources of Information on Breast Implant-Associated Anaplastic Large Cell Lymphoma and Breast Implant Illness. *Aesthetic Plast. Surg.* **2023**, *48*, 590–607. [CrossRef] [PubMed]
35. Rahsepar, A.A.; Tavakoli, N.; Kim, G.H.J.; Hassani, C.; Abtin, F.; Bedayat, A. How AI Responds to Common Lung Cancer Questions: ChatGPT versus Google Bard. *Radiology* **2023**, *307*, e230922. [CrossRef] [PubMed]
36. Emile, S.H.; Horesh, N.; Freund, M.; Pellino, G.; Oliveira, L.; Wignakumar, A.; Wexner, S.D. How Appropriate Are Answers of Online Chat-Based Artificial Intelligence (ChatGPT) to Common Questions on Colon Cancer? *Surgery* **2023**, *174*, 1273–1275. [CrossRef] [PubMed]
37. Moazzam, Z.; Cloyd, J.; Lima, H.A.; Pawlik, T.M. Quality of ChatGPT Responses to Questions Related to Pancreatic Cancer and Its Surgical Care. *Ann. Surg. Oncol.* **2023**, *30*, 6284–6286. [CrossRef] [PubMed]
38. Hermann, C.E.; Patel, J.M.; Boyd, L.; Growdon, W.B.; Aviki, E.; Stasenko, M. Let’s Chat about Cervical Cancer: Assessing the Accuracy of ChatGPT Responses to Cervical Cancer Questions. *Gynecol. Oncol.* **2023**, *179*, 164–168. [CrossRef] [PubMed]
39. Chow, J.C.L.; Wong, V.; Sanders, L.; Li, K. Developing an AI-Assisted Educational Chatbot for Radiotherapy Using the IBM Watson Assistant Platform. *Healthcare* **2023**, *11*, 2417. [CrossRef] [PubMed]
40. Stokel-Walker, C. AI Bot ChatGPT Writes Smart Essays—Should Professors Worry? *Nature*, 9 December 2022; d41586-022-04397-7. [CrossRef]
41. Hopkins, A.M.; Logan, J.M.; Kichenadasse, G.; Sorich, M.J. Artificial Intelligence Chatbots Will Revolutionize How Cancer Patients Access Information: ChatGPT Represents a Paradigm-Shift. *JNCI Cancer Spectr.* **2023**, *7*, pkad010. [CrossRef] [PubMed]

42. Nedbal, C.; Naik, N.; Castellani, D.; Gauhar, V.; Geraghty, R.; Somani, B.K. ChatGPT in Urology Practice: Revolutionizing Efficiency and Patient Care with Generative Artificial Intelligence. *Curr. Opin. Urol.* **2024**, *34*, 98–104. [[CrossRef](#)] [[PubMed](#)]
43. Dahmen, J.; Kayaalp, M.E.; Ollivier, M.; Pareek, A.; Hirschmann, M.T.; Karlsson, J.; Winkler, P.W. Artificial Intelligence Bot ChatGPT in Medical Research: The Potential Game Changer as a Double-Edged Sword. *Knee Surg. Sports Traumatol. Arthrosc.* **2023**, *31*, 1187–1189. [[CrossRef](#)] [[PubMed](#)]
44. Whiles, B.B.; Bird, V.G.; Canales, B.K.; DiBianco, J.M.; Terry, R.S. Caution! AI Bot Has Entered the Patient Chat: ChatGPT Has Limitations in Providing Accurate Urologic Healthcare Advice. *Urology* **2023**, *180*, 278–284. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Development of AI-generated medical responses using the ChatGPT for cancer patients



Jae-woo Lee^{a,b,1}, In-Sang Yoo^{c,h,1}, Ji-Hye Kim^c, Won Tae Kim^{d,e}, Hyun Jeong Jeon^{f,g}, Hyo-Sun Yoo^a, Jae Gwang Shin^c, Geun-Hyeong Kim^c, ShinJi Hwang^c, Seung Park^{c,h,2}, Yong-June Kim^{d,e,2,*}

^a Department of Family Medicine, Chungbuk National University Hospital, Cheongju, Republic of Korea

^b Department of Family Medicine, Chungbuk National University College of Medicine, Cheongju, Republic of Korea

^c Department of Biomedical Engineering, Chungbuk National University Hospital, Cheongju, Republic of Korea

^d Department of Urology, Chungbuk National University Hospital, Cheongju, Republic of Korea

^e Department of Urology, Chungbuk National University College of Medicine, 1 Chungdae-ro, Seowon-gu, Cheongju, Chungcheongbuk-do 28644, Republic of Korea

^f Department of Internal Medicine, Chungbuk National University Hospital, Cheongju, Republic of Korea

^g Department of Internal Medicine, College of Medicine, Chungbuk National University, Cheongju, Republic of Korea

^h Department of Medicine, Chungbuk National University College of Medicine, Cheongju, Republic of Korea

ARTICLE INFO

Keywords:

Chatbot
AI
Cancer patients
ChatGPT
Natural language processing models

ABSTRACT

Background and objective: To develop a healthcare chatbot service (AI-guided bot) that conducts real-time conversations using large language models to provide accurate health information to patients.

Methods: To provide accurate and specialized medical responses, we integrated several cancer practice guidelines. The size of the integrated meta-dataset was 1.17 million tokens. The integrated and classified metadata were extracted, transformed into text, segmented to specific character lengths, and vectorized using the embedding model. The AI-guide bot was implemented using Python 3.9. To enhance the scalability and incorporate the integrated dataset, we combined the AI-guide bot with OpenAI and the LangChain framework. To generate user-friendly conversations, a language model was developed based on Chat-Generative Pretrained Transformer (ChatGPT), an interactive conversational chatbot powered by GPT-3.5. The AI-guide bot was implemented using ChatGPT3.5 from Sep. 2023 to Jan. 2024.

Results: The AI-guide bot allowed users to select their desired cancer type and language for conversational interactions. The AI-guided bot was designed to expand its capabilities to encompass multiple major cancer types.

The performance of the AI-guide bot responses was 90.98 ± 4.02 (obtained by summing up the Likert scores). **Conclusions:** The AI-guide bot can provide medical information quickly and accurately to patients with cancer who are concerned about their health.

1. Introduction

Cancer is one of the leading causes of death worldwide; it accounted for over 19.3 million new diagnoses and 10 million deaths in 2020 [1]. According to the CONCORD-3 reports, the five-year net survival trends have increased for most cancers, including the cases associated with some of the more lethal cancers [2]. Although there are differences among cancer sites and countries, cancer survival rates are high and continuously increasing [2]. By 2030, approximately 21.6 million

patients per year worldwide are expected to be diagnosed with cancer [3]. All cancer patients need prompt diagnosis and optimal treatment to improve survival. Therefore, the importance of an effective healthcare system for the treatment and care of patients with cancer is increasing. Cancer patients have several concerns about physical, practical, and emotional problems, but they do not receive appropriate information to manage these concerns [4]. In particular, patients often receive misinformation about their health from Internet sources, such as Google or YouTube [5–7]. In previous studies, [6], it was reported that harmful

* Corresponding author.

E-mail address: urokyj@gmail.com (Y.-J. Kim).

¹ These authors contributed equally to the manuscript.

² These authors contributed equally to this work.

<https://doi.org/10.1016/j.cmpb.2024.108302>

Received 22 January 2024; Received in revised form 28 May 2024; Accepted 22 June 2024

Available online 26 June 2024

0169-2607/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

misinformation about cancer was being disseminated through YouTube and that misinformation was being indiscriminately uploaded to the Internet not only by nonprofessional groups but also by professional groups.

Conversely, natural language processing (NLP) models can be applied to enhance access to medical information for patients [8]. Large language models (LLMs) are NLP tools that can understand and generate human-like text, such as the Chat-Generative Pretrained Transformer (ChatGPT) [9]. ChatGPT is a conversational chatbot based on GPT-3.5 with 175 billion parameters trained on a wide range of sources, such as books and articles [10]. This chatbot was fine-tuned for conversational tasks using reinforcement learning with human feedback. This training enables it to handle the complexity of user intent and respond effectively [11]. Conventional LLMs are simpler models predicting sequences based on fixed windows of words, whereas ChatGPT utilizes vast datasets to generate human-like conversations [9]. By using ChatGPT to develop chatbots for cancer patients, it can be expected to improve dramatically the accessibility of medical information. In addition, effective health-information delivery can be an alternative solution to the unsatisfied health information needs of patients.

In this study, we aimed to develop a healthcare chatbot service (hereafter referred to as the AI-guide bot) that conducts real-time conversations using an LLM model designed by a cancer-related expert group to provide accurate health information to patients. This service is expected to provide assistance to address the unmet needs of patients with cancer.

2. Methods

2.1. Data source

To develop a real-time chatbot in which AI answers medical questions for cancer patients, we first constructed a specialized medical knowledge database related to cancer. To provide accurate and specialized medical responses, cancer experts participated in the construction of the dataset using accredited medical guidelines. For prostate cancer, we integrated prostate cancer practice guidelines from the Korean Prostate Society, [12], accredited health guidelines for patients, such as the National Comprehensive Cancer Network, American Society of Clinical Oncology, and European Association of Urology Guidelines, [13–18], urology book, [19,20], and instructional videos [21]. The total size of the integrated meta-dataset was 1.17 million tokens.

The integrated meta-dataset was divided into the following specific categories: definition, epidemiology, causes, symptoms, diagnosis, staging, treatment, prognosis, and lifestyle. The categorized meta-dataset was organized by considering actual questions posed by patients and nonspecialized medical providers as well as questions raised by cancer experts (Supplementary Fig. 1). The integrated and classified metadata were extracted, transformed into text, segmented to a specific character length, and vectorized using the embedding model provided by OpenAI [22].

This study was approved by the Institutional Review Board of the Chungbuk National University Hospital (CBNUH-04-020) and adhered to the guidelines of the Declaration of Helsinki (1975). The Ethics Committee waived the need for informed consent because the data from this study were anonymized at all stages.

2.2. Development process

The AI-guide bot was implemented using Python 3.9. To enhance the scalability and incorporate the integrated dataset, we combined the AI-guide bot with OpenAI [23] and the LangChain framework [24]. To generate user-friendly conversations, a language model was developed based on ChatGPT, an interactive conversational chatbot powered by GPT-3.5. The AI-guide bot was implemented using ChatGPT3.5 from Sep. 2023 to Jan. 2024. We adopted the GPT 3.5 version as our baseline

rather than GPT 4.0 owing to its cost-effectiveness. The collected data, described in the “Data source” section, was vectorized using the OpenAI embedding model and stored in a database using the FAISS Python library, a vector index (Fig. 1). When a user submits a question, the AI-Chatbot evaluates whether the query contains cancer-related information in the database. We used a fixed prompt for this study and deliberately set the temperature parameter to a relatively low value of 0.1 (default = 0.7). This decision was based on our objective to generate responses that are consistent and closely aligned with the intent of the questions posed. Additionally, while modifying the GPT prompts, we aimed to ensure consistent responses by repeating the same questions 10 times. By adopting this setting, we observed that the model predominantly produced answers that were coherent and relevant to the questions asked. If the question is unrelated to the designated type of cancer, the AI-Chatbot is configured to respond that it could not provide an answer because of the irrelevance of its content as the database contains medically specialized terminologies, the AI-Chatbot interprets the user's intent and formulates an appropriate keyword. The refined question then undergoes vectorization, and its similarity is assessed in relation to each entry in the database. To ensure that the GPT generated a response using the pertinent information, we incorporated the top four similar entries into the GPT prompt.

To facilitate accurate communication of medical information, when the similarity score of the retrieved response vector index was below 0.4, the query was output as unrelated to the specific cancer type or owing to a lack of information. This prompts users to rephrase their questions more accurately. When the similarity score of the retrieved response vector index was 0.4 or higher, a prompt was constructed using the user's question, selected vector index with the highest similarity, and the predefined instructions for structuring the output sentence. For questions unrelated to health, we confirmed that the model tended to reject generating answers. This prompt was then delivered to the output server of the AI-guide bot. The fixed directives were designed in a simple and clear format to induce familiar and easy conversations with the user. In cases wherein the responses generated by the AI-guide bot for the output were not very readable, they were corrected and output in user-friendly terms. This entire output process was recorded as part of the conversation history between the user and the AI-guide bot, thus making it accessible for future prompts (Fig. 1, Supplementary Fig. 2).

2.3. Performance evaluation process

To evaluate the performance of the AI-guide bot, we constructed a cancer-related question dataset and assessed whether the AI-guide bot provided correct and accurate medical responses. A total of 100 questions were validated by multiple oncology experts to ensure clear and uncontroversial answers in line with the latest medical guidelines. All questions were selected to represent expert knowledge in oncology (Supplementary Table 1).

The evaluation of the AI-guide bot responses to cancer-related questions was conducted by specialists (Urology, Family Medicine, Endocrinology, and Hematology–Oncology) who were not involved in developing the cancer-related question dataset. The response evaluation focused on whether the AI-guide bot clearly understood the patient's question (comprehensibility), accuracy of the response content (accuracy), and whether understanding was easy (readability). Comprehensibility was based on how well the AI-guide bot understood the intent of the questions, determining whether the guide bot effectively set the direction for its responses by analyzing the questions entered into the natural language. A Likert score for comprehensibility was assigned to each question on a scale of 10 units (between 10 and 30). Example (10) indicates that the intent of the question and the response are entirely unrelated. Example (20) implies that unrelated content is included; however, the response is relevant to the intent of the question. Example (30) was assigned when a response was considered appropriate for the question's intent. Content accuracy was assessed based on the accuracy

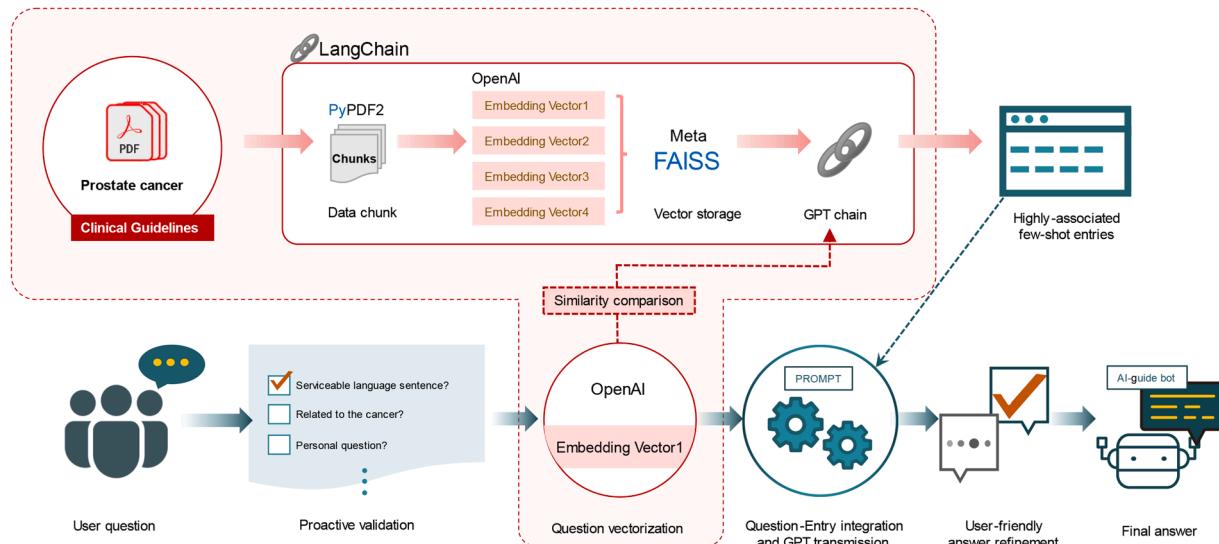


Fig. 1. Development process of artificial intelligence (AI)-guide bot for cancer patients.

of the AI-generated medical responses, regardless of the intent of the question. Content accuracy was established to evaluate the reliability of the responses because a guidebot needs to provide medical-related answers. A Likert score (ranging from 10 to 40 on a scale of 10 units) was assigned to each question for content accuracy. Example (10) indicates that there was no accuracy in the response regardless of the question's intent. Example (20) implies that some facts are included but most information is false. Example (30) is assigned when false odds are included; however, most of the information is accurate. Example (40) was awarded when the response was considered to provide accurate information. Readability evaluates the comprehension difficulty and naturalness of sentences irrespective of the question's intent and accuracy. Sentence completeness is designed to assess the complexity of sentences and whether they are difficult to understand, as the guide bot is expected to communicate with the general public. Unlike the previous evaluation criteria, the Likert scores were set on a narrower scale because readability is subjective. Additionally, readability solely evaluates sentence completeness, which is unrelated to question relevance or content accuracy. Each question was assigned a Likert score (which ranged from 10 to 30 on a 10-unit scale). Example (10) implies that the use of medical terminology is convoluted, and grammar is incorrect, thus making it difficult to understand. Example (20) suggests that grammar is correct; however, the excessive use of medical terms makes it challenging. Example (30) is assigned when the sentence structure is appropriate and is not considered difficult for adults to understand.

The performance of the AI-guided bot responses was assessed by summing their Likert scores. Responses with a total Likert score of 30 or lower were classified as “inappropriate,” those with total scores of 70 or lower were considered to include some “minor errors or they were incomplete,” and responses with scores above 70 were classified as “appropriate” for the questions [25] (Table 1).

As an additional performance evaluation method, considering that the AI-guide bot user was a cancer patient, performance evaluation was also conducted on a group of 50 non-medical experts. Since this was a non-expert group, accuracy was excluded and a performance evaluation comparison was conducted using ChatGPT in the areas of comprehensibility and readability.

3. Results

The AI-guide bot allows users to select their desired cancer type and language for conversational interactions. Both English and Korean languages provided language support for the study. Users have the

Table 1

Explanation and performance of artificial intelligence (AI)-guide bot for cancer patients.

Criteria	Explanation	Performance (Likert sum average)
Comprehensibility	10 The question's intent and the response are entirely unrelated	28.28 ± 0.38
	20 While some unrelated content is mixed, the response is relevant to the intent of the question	
	30 The response is considered appropriate to the intent of the question	
Accuracy	10 No accurate information in the response	34.17 ± 2.91
	20 Some facts are included, but most of the information is false	
	30 Some falsehoods are included, but most of the information is accurate	
	40 The response is considered to provide entirely accurate information	
Readability	10 The response is difficult owing to the disorganized use of technical terminology and lack of proper grammar	28.53 ± 1.24
	20 The grammar is clean, but an excessive use of medical terms makes it challenging	
	30 Sentence structure is appropriate and it is not considered difficult for adults to understand	
Total	100	90.98 ± 4.02

flexibility to switch to different cancer types or languages during conversations with the AI-guided bot. An AI-generated Chatbot was designed to expand its capabilities to encompass multiple major cancer types. The initial model was developed with a strong focus on prostate cancer, benefiting from established clinical guidelines. Fig. 2 shows a screenshot of the actual output results of the AI-guide bot. When a user asks a question related to cancer, the AI-guide bot understands the question and outputs an appropriate and accurate answer from the medical guideline metadata. An example of a slightly more complex question is given here; “I was diagnosed with stage 1 prostate cancer.

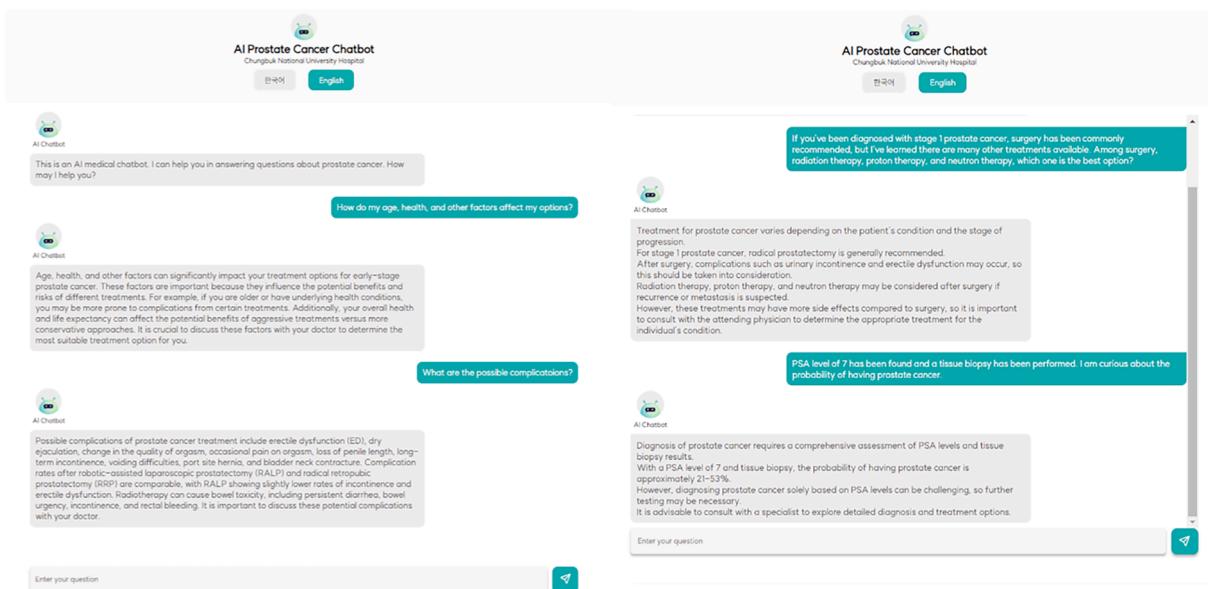


Fig. 2. Output screenshot of AI-Guide bot for cancer patients.

The hospital recommended surgery, but I've heard other treatments are available. What's the best option between surgery, radiation therapy, proton therapy, and neutron therapy?" For this question (supplementary Table 1, #88. Question), the AI-guide bot's automatic medical answer output was as follows; "Treatment for prostate cancer varies depending on the patient's condition and the stage of progression. For stage 1 prostate cancer, radical prostatectomy is generally recommended. After surgery, complications such as urinary incontinence and erectile dysfunction may occur, so this should be taken into consideration. Radiation therapy, proton therapy, and neutron therapy may be considered after surgery if recurrence or metastasis is suspected. However, these treatments may have more side effects compared to surgery, so it is important to consult with the attending physician to determine the appropriate treatment for the individual's condition". In the case of prostate cancer, the PSA level is often checked during health checkups, and patients are curious about the meaning of the results of this test. When the PSA-related questions: "My PSA is seven, and I had a tissue biopsy. I'm wondering about the likelihood of having prostate cancer" (Supplementary Table 1. #89. Question) were entered into the AI-guide bot, and the output was "Diagnosis of prostate cancer requires a comprehensive assessment of PSA levels and tissue biopsy results. With a PSA level of 7 and a tissue biopsy, the probability of having prostate cancer is approximately 21–53 %. However, diagnosing prostate cancer solely based on PSA levels can be challenging, so further testing may be necessary. The answer is generated as "It is advisable to consult with a specialist to explore detailed diagnosis and treatment options." This answer provides information such as the meaning of PSA level and the incidence of prostate cancer that roughly corresponds to this level (Fig. 2).

Table 1 presents the explanation and performance evaluation results of medical experts group for the AI-guide bot. Understanding the user's intention to ask questions was assigned a score of 28.28 ± 0.38 , accuracy of content was assigned to a score of 34.17 ± 2.91 , and readability was assigned to a score of 28.53 ± 1.24 . The performance of the AI-guide bot responses was 90.98 ± 4.02 (obtained by summing up the Likert scores). From the performance evaluation results conducted by a non-expert group, the AI-guide bot achieved statistically higher results than ChatGPT in comprehension and readability (Table 2).

Table 2

Performance comparison with ChatGPT in a non-expert group for artificial intelligence (AI)-guide bot.

	AI-guide bot Mean \pm S.D	ChatGPT Mean \pm S.D	P-value
Comprehensibility	28.4 ± 0.38	26.1 ± 0.81	<0.0001
Readability	28.3 ± 0.39	25.6 ± 0.84	<0.0001

Abbreviations: S.D, Standard deviation

The P-value was calculated using the t-test between the two groups.

4. Discussion

4.1. Main findings

In this study, a ChatGPT-based health consultation chatbot was developed for cancer patients using an LLM model. This healthcare chatbot program yielded promising results and can be used for cancer patients to understand the intent of the question, accuracy of the answers, and readability. In addition, through the collaboration of cancer-related experts, the accuracy of the answers was reverified so that false information was not provided to patients. We expect the chatbot program to become a compass for health promotion among patients with cancer.

4.2. AI-generated medical responses

Previously, most NLP models were trained using supervised learning and could be used for trained tasks only [26]. However, in 2017, with the introduction of a self-attention network architecture known as a transformer, [27]. ChatGPT was able to apply pretrained language representations to downstream tasks through a combination of unsupervised pretraining and supervised fine-tuning. The strength of ChatGPT lies in its NLP capabilities, which are based on GPT-3.5. With this feature, ChatGPT can extract information from unstructured medical information data sources, such as medical guidelines and medical textbooks, identify patterns, such as specific symptoms, and generate output results from them [28]. ChatGPT has recently been upgraded to include new features and improvements. The new LLM (GPT-4) can process up to 25,000 words and boost creativity and visual input ahead of its predecessor, GPT-3.5, as well as provide longer contextual

memory. According to the internal evaluation, ChatGPT has now become a safer and more accurate tool for tasks, thus providing 40 % more reliable responses [28].

In this study, the developed AI-guide bot utilized AI and natural language processing, similar to other traditional medical chatbots, [29] to understand user intentions and provide appropriate responses tailored to user requirements by retrieving information from medical datasets. Additionally, AI-guide bot can be further improved through user feedback, enabling the provision of knowledge pertaining to various clinical scenarios of the patients' clinical needs. This ongoing process empowers us to customize the performance of AI-guide bots according to individual patient requirements.

In contrast, ChatGPT raises concerns about the accuracy and reliability of the medical information it provides, as it draws information from the internet [28]. To address this issue, medical experts provide meticulous training to the AI-guide bot, thus preventing the dissemination of inaccurate information to patients. In this study, the performance of the AI-guided bot was evaluated using three criteria: comprehensibility, accuracy, and readability. The average total Likert score was 90.98, which surpasses the acceptable threshold of 70. A previous study examined the information quality and adequacy of the ChatGPT output for urological patients, and only 52 % of all responses reported that it was adequate [30]. The performance results of the AI-guide bot developed in this study demonstrate its potential for ongoing development owing to its adaptability through database modifications and additional AI training by medical experts. This adaptability constitutes a considerable advantage for future programs.

4.3. Unmet needs of cancer patients

The provision of medical information to cancer patients can be expected to yield several benefits, including increased patient participation in decision-making and enhanced satisfaction with medical treatment [31,32]. Moreover, it may improve coping abilities and reduce anxiety and worries related to cancer diagnosis, treatment, and post-treatment phases [33]. This patient information has been recognized as an effective individual means to cope with the quality of life impairment associated with cancer diagnosis and treatment [34]. Furthermore, effective communication between physicians and patients is crucial for optimal health outcomes in the medical field [34]. It is known that it is essential to ensure that patients receive and understand appropriate medical information during their treatment process to ensure the provision of quality cancer treatment [31–34].

Medical chatbots offer an effective communication method with ChatGPT and provide a user-friendly and concise conversational approach to information delivery [9]. Cancer patients sometimes have unmet medical information needs, such as on the Internet or YouTube [5–7]. In this regard, for cancer patients in need of accurate medical information, the AI-guided bot can serve as a suitable gateway for effective communication and as a means to address unmet needs. Because the information quality and adequacy of ChatGPT output are not accurate in urological patients, [30] to compensate for information adequacy, we trained several guidelines to improve the quality and readability of the information.

4.4. Role of AI-guide bot in AI-driven infodemic

A potential concern of LLMs, such as ChatGPT, is the submissiveness of users in the following instructions. Despite the restrictions imposed by developers, LLMs can be easily used to generate text on controversial topics, including misinformation [9]. LLMs' ability to generate text similar to human-written texts can be used to create fake news or fabricated or misleading content without the reader realizing that the text was generated by AI [35–37]. The spread of misinformation through fast-disseminating channels, such as mass media and social networks, can lead to a phenomenon known as the infodemic [38]. Infodemic can

affect medical decision-making regarding treatment or preventive measures [39]. Social media platforms and other media outlets can quickly disseminate misinformation and generate an abundant of health-related information inaccurately. Given the nature of social media content proliferation, LLMs can be used to create content for a specific target population and promote the spread of misinformation. With these characteristics of LLM, the concept of infodemic was presented as the concept of "AI-driven infodemic" [9].

Medical information provided to patients must be accurate. Furthermore, it should not infringe on the authority of physicians' medical decisions. It may be prudent not to provide any medical information to patients rather than offering false or unverified medical information. This is particularly true for cancer, which is a pathologically complex clinical condition. Failure to provide cancer patients with accurate information can exacerbate confusion rather than alleviate it. While ChatGPT has passed the United States Medical Licensing Exam, [40] it cannot replace human interactions in the pursuit of more inclusive healthcare [9,41]. Additionally, in AI-driven infodemic situations, there is a risk that ChatGPT may provide incorrect answers. If indiscriminate medical information overflows and AI provides AI-generated answers using incorrect information from the Internet, it can confuse the treatments of cancer patients. In this respect, an AI-guide bot was developed to refine inaccurate medical information overflowing on the Internet and to provide accurate medical information to cancer patients.

4.5. Strengths and limitations

In the course of formulating an AI-driven medical guidance chatbot, designated as an AI-guide bot, via integration of the GPT API and Langchain technology, several limitations need to be considered. First, user-derived feedback is lacking. The AI-guide bot must emphasize the importance of garnering authentic user responses. Augmentation of the bot's precision and functional efficacy mandates a systematic aggregation of such feedback, accompanied by a rigorous analysis. Second, there are inherent limitations to the FAISS similarity-check algorithm. This limitation manifests itself in the obstacles concerning the apprehension and application of medical terminological synonyms. Periodic curation and reinforcement of a medical synonym repository are required to improve the accuracy of the query resolution of the AI-guide bot. Third, we aimed to ensure consistent responses by repeating the same question 10 times while fixing the GPT prompts. This approach yielded consistent answers in the vast majority of cases, with only approximately 2–3 % of the responses deviating from the intended question. Finally, the AI-guided bot may have concerns related to data privacy and infringement. Utilization of the GPT API engenders existing anxieties about data privacy and possible violations thereof. Consequently, it is critical to incorporate protocols for personal data detection and obfuscation in the initial phases of development to safeguard personal information and ensure regulatory compliance. By recognizing and ameliorating these impediments, we can further improve the service quality and operational efficacy of the AI-guide bot.

Despite these limitations, this study had several strengths. The greatest advantage is that the chatbot systems used in this study offer significant advantages in improving the quality of medical care. While the AI-guide bot cannot replace medical professionals, it can assist patients in understanding their medical conditions, the importance of their ongoing diagnostic tests, and disease medical treatments by doctors, thus empowering them with self-awareness. The accurate diagnosis and treatment provided by a physician are essential for treating a patient, but equally important is how well patients adhere to treatment of their disease [34]. The AI-guide bot can serve as a supplementary tool for patient healthcare. In real medical care settings, there are instances in which doctors may lack the time to provide thorough explanations to patients, or even if they do, patients may have difficulty comprehending the medical information. In such situations, an AI-guide bot can play a significant role in a patient's medical journey.

The second strength of this study is that the AI-guided bot can provide precise medical information supported by the involvement of clinicians, and the data can be continually improved through learning. Notably, our AI-guide bot has demonstrated robust outcomes regarding AI hallucinations due to its design to provide information solely from within its database. Furthermore, while currently being developed for prostate cancers, another advantage is the possibility of expanding the scope of provision for cancer patients in the future. Lastly, since there are very few chatbot systems in the world for cancer patients in which AI provides medical information that is user-friendly and appropriate for the patient's unmet medical information needs, the development of an AI-guide bot can be considered valuable.

In conclusion, if the healthcare chatbot program for cancer patients developed in this study is used, it will be possible to provide medical information quickly and accurately to patients who are concerned about their health. More follow-up studies are needed to provide accurate information. However, it is believed that the developed chatbot program could help improve the health of cancer patients.

Funding

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (Grant No. HR22C1302).

Data availability

Data from this study may be shared upon reasonable request from the corresponding author.

CRediT authorship contribution statement

Jae-woo Lee: Writing – review & editing, Writing – original draft, Visualization, Validation, Investigation. **In-Sang Yoo:** Methodology, Formal analysis, Data curation. **Ji-Hye Kim:** Software, Resources, Methodology, Formal analysis, Data curation. **Won Tae Kim:** Validation, Supervision. **Hyun Jeong Jeon:** Writing – review & editing, Validation, Supervision. **Hyo-Sun Yoo:** Validation, Supervision, Project administration, Investigation. **Jae Gwang Shin:** Project administration, Methodology, Formal analysis, Data curation. **Geun-Hyeong Kim:** Methodology, Formal analysis, Data curation. **ShinJi Hwang:** Software, Project administration, Methodology, Formal analysis, Data curation. **Seung Park:** Writing – review & editing, Validation, Supervision, Software, Data curation, Conceptualization. **Yong-June Kim:** Writing – review & editing, Validation, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors have no potential conflicts of interest to disclose.

Acknowledgment

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (Grant No. HR22C1302).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cmpb.2024.108302.

References

- [1] B.S. Chhikara, K. Parang, Global Cancer Statistics 2022: the trends projection analysis, *Chem. Biol. Lett.* 10 (2023) 451.
- [2] C. Allemani, T. Matsuda, V. Di Carlo, R. Harewood, M. Matz, M. Nikšić, A. Bonaventure, M. Valkov, C.J. Johnson, J. Estève, O.J. Ogundbiyi, E.S.G. Azevedo, W.Q. Chen, S. Eser, G. Engholm, C.A. Stiller, A. Monnereau, R.R. Woods, O. Visser, G.H. Lim, J. Aitken, H.K. Weir, M.P. Coleman, Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries, *Lancet* 391 (2018) 1023–1075.
- [3] World Health Organization, Cancer Prevention and Control in the Context of an Integrated approach: Report By the Secretariat [Internet], WHO, Geneva, 2016 [cited 2019 Apr 29].
- [4] M.I. Fitch, I. Nicoll, G. Lockwood, Cancer survivor's perspectives on the major challenge in the transition to survivorship, *Patient Educ. Couns.* 103 (2020) 2361–2367.
- [5] M. Wasserman, N.N. Baxter, B. Rosen, M. Burnstein, A.L. Halverson, Systematic review of internet patient information on colorectal cancer surgery, *Dis. Colon Rectum* 57 (2014) 64–69.
- [6] H. Park, E. Kang, Y. Kim, H. Ju, Analysis of the spread of misinformation about Lung Cancer on youtube: based on source of information, *Korean J. Fam. Pract.* 13 (2023) 152–158.
- [7] A.N. Sahin, A.S. Sahin, F. Schwenter, H. Sebagang, Youtube videos as a source of information on colorectal cancer: what do our patients learn? *J. Cancer Educ.* 34 (2019) 1160–1166.
- [8] S. Locke, A. Bashall, S. Al-Adely, J. Moore, A. Wilson, G.B. Kitchen, Natural language processing in medicine: a review, *Trends Anaesth. Crit. Care* 38 (2021) 4–9.
- [9] L. De Angelis, F. Baglivo, G. Arzilli, G.P. Privitera, P. Ferragina, A.E. Tozzi, C. Rizzo, ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health, *Front. Public Health* 11 (2023) 1166120.
- [10] T. Brown, B. Mann, N. Ryder, M. Subbiah, J.D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, Language models are few-shot learners, *Adv. Neural Inf. Process. Syst.* 33 (2020) 1877–1901.
- [11] R.S. Goodman, J.R. Patrinely, C.A. Stone Jr., E. Zimmerman, R.R. Donald, S. Chang, S.T. Berkowitz, A.P. Finn, E. Jahangir, E.A. Scoville, T.S. Reese, D. L. Friedman, J.A. Bastarache, Y.F. van der Heijden, J.J. Wright, F. Ye, N. Carter, M. R. Alexander, J.H. Choe, C.A. Chastain, J.A. Zic, S.N. Horst, I. Turker, R. Agarwal, E. Osmundson, K. Idrees, C.M. Kiernan, C. Padmanabhan, C.E. Bailey, C. E. Schlegel, L.B. Chambliss, M.K. Gibson, T.J. Osterman, L.E. Wheless, D. B. Johnson, Accuracy and reliability of chatbot responses to physician questions, *JAMA Netw. Open* 6 (2023) e2336483.
- [12] The Korean prostate society. Prostate Cancer Guidelines, 2021.
- [13] The Korean urological oncology society. Guidance for Prostate Cancer Patients. 2019.
- [14] The Korean Urological Association. The medical treatment of prostate hyperplasia. 2020.
- [15] The Korean Cancer Association. Nutrition and physical activity guidelines for cancer patients. 2019.
- [16] K.A. Moses, P.C. Sprenkle, C. Bahler, G. Box, S.V. Carlsson, W.J. Catalona, D. M. Dahl, M. Dall'Era, J.W. Davis, B.F. Drake, NCCN guidelines® insights: prostate cancer early detection, version 1.2023: featured updates to the NCCN guidelines, *J. Natl. Compr. Cancer Netw.* 21 (2023) 236–246.
- [17] K.S. Virgo, R.B. Rumble, R. de Wit, D.S. Mendelson, T.J. Smith, M.-E. Taplin, J. L. Wade III, C.L. Bennett, H.L. Scher, P.L. Nguyen, Initial management of noncastrate advanced, recurrent, or metastatic prostate cancer: ASCO guideline update, *J. Clin. Oncol.* 39 (2021) 1274–1305.
- [18] N. Mottet, R.C.N. van den Berg, E. Briers, T. Van den Broeck, M.G. Cumberbatch, M. De Santis, S. Fanti, N. Fossati, G. Gandaglia, S. Gillessen, N. Grivas, J. Grummert, A.M. Henry, T.H. van der Kwast, T.B. Lam, M. Lardas, M. Liew, M.D. Mason, L. Moris, D.E. Oprea-Lager, H.G. van der Poel, O. Rouvière, I.G. Schoots, D. Tilki, T. Wiegel, P.M. Willemse, P. Cornford, EAU-EANM-ESTRO-ESUR-SIOP guidelines on prostate cancer-2020 update. part 1: screening, diagnosis, and local treatment with curative intent, *Eur Urol* 79 (2021) 243–262.
- [19] The Korean Prostate Society, Getting to Know the Prostate, 4th ed., Ilchokak, 2023.
- [20] The Korean Urological Association, Urology 6th Edition, Ilchokak, 2019.
- [21] HealthBreeze, Medical Animation <https://www.healthbreeze.com/>.
- [22] M. Schwarz, K. Chapman, B. Häussler, Multilingual medical entity recognition and cross-lingual zero-shot linking with facebook AI similarity search, (2022).
- [23] OpenAI, Introducing ChatGPT. 2024 <https://openai.com/blog/chatgpt/>.
- [24] LangChain 2024 https://python.langchain.com/docs/get_started/introduction.
- [25] J. Rudolph, S. Tan, S. Tan, War of the chatbots: bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education, *J. Appl. Learn. Teach.* 6 (2023).
- [26] R. Shams, Semi-supervised classification for natural language processing, arXiv preprint arXiv:1409.7612, (2014).
- [27] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, *Adv. Neural Inf. Process. Syst.* (2017) 30.
- [28] J.C. Chow, L. Sanders, K. Li, Impact of ChatGPT on medical chatbots as a disruptive technology, *Front. Artif. Intell.* 6 (2023) 1166014.
- [29] J.C.L. Chow, L. Sanders, K. Li, Design of an educational chatbot using artificial intelligence in radiotherapy, *AI* 4 (2023) 319–332.
- [30] A. Coccia, M. Pezzoli, M. Lo Re, G.I. Russo, M.G. Asmundo, M. Fode, G. Cacciamani, S. Cimino, A. Minervini, E. Durukan, Quality of information and appropriateness of ChatGPT outputs for urology patients, *Prostate Cancer Prostatic Dis.* (2023).

- [31] K.A. Luker, K. Beaver, S.J. Leinster, R.G. Owens, L.F. Degner, J.A. Sloan, The information needs of women newly diagnosed with breast cancer, *J. Adv. Nurs.* 22 (1995) 134–141.
- [32] M. Cawley, J. Kostic, C. Cappello, Informational and psychosocial needs of women choosing conservative surgery/primary radiation for early stage breast cancer, *Cancer Nurs.* 13 (1990) 90–94.
- [33] D.E. Harrison-Woermke, J.E. Graydon, Perceived informational needs of breast cancer patients receiving radiation therapy after excisional biopsy and axillary node dissection, *Cancer Nurs.* 16 (1993) 449–455.
- [34] L.J.F. Rutten, N.K. Arora, A.D. Bakos, N. Aziz, J. Rowland, Information needs and sources of information among cancer patients: a systematic review of research (1980–2003), *Patient Educ. Couns.* 57 (2005) 250–261.
- [35] R.F. Sear, R. Leahy, N.J. Restrepo, Y. Lupu, N.F. Johnson, Machine learning language models: achilles heel for social media platforms and a possible solution, *Adv. Artif. Intell. Mach. Learn.* 1 (2021) 191–202.
- [36] A. Gupta, A. Singhal, A. Mahajan, A. Jolly, S. Kumar, Empirical framework for automatic detection of neural and human authored fake news, in: Proceedings of the 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), IEEE, 2022, pp. 1625–1633.
- [37] K. McGuffie, A. Newhouse, The radicalization risks of GPT-3 and advanced neural language models, arXiv preprint arXiv:2009.06807, (2020).
- [38] The lancet infectious D. The COVID-19 infodemic, *Lancet Infect. Dis.* 20 (2020) 875.
- [39] S.C. Briand, M. Cinelli, T. Nguyen, R. Lewis, D. Prybylski, C.M. Valensise, V. Colizza, A.E. Tozzi, N. Perra, A. Baronchelli, Infodemics: a new challenge for public health, *Cell* 184 (2021) 6010–6014.
- [40] T.H. Kung, M. Cheatham, A. Medenilla, C. Sillos, L. De Leon, C. Elepaño, M. Madriaga, R. Aggabao, G. Diaz-Candido, J. Maningo, V. Tseng, Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models, *PLOS Digit. Health* 2 (2023) e0000198.
- [41] J.E.H. Brown, J. Halpern, AI chatbots cannot replace human interactions in the pursuit of more inclusive mental healthcare, *SSM Mental Health* 1 (2021) 100017.

Accepted Manuscript

Review

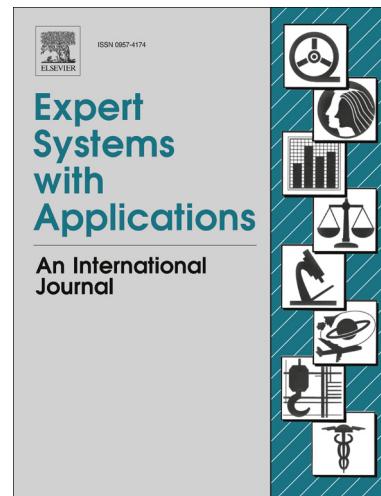
Image Based Computer Aided Diagnosis System for Cancer Detection

Howard Lee, Yi-Ping Phoebe Chen

PII: S0957-4174(15)00098-6

DOI: <http://dx.doi.org/10.1016/j.eswa.2015.02.005>

Reference: ESWA 9855



To appear in: *Expert Systems with Applications*

Please cite this article as: Lee, H., Chen, Y.P., Image Based Computer Aided Diagnosis System for Cancer Detection, *Expert Systems with Applications* (2015), doi: <http://dx.doi.org/10.1016/j.eswa.2015.02.005>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Image Based Computer Aided Diagnosis System for Cancer Detection

Howard Lee, Yi-Ping Phoebe Chen*,

Department of Computer Science and Computer Engineering, La Trobe University, Melbourne, Australia

H17lee@stuents.latrobe.edu.au; phoebe.chen@latrobe.edu.au

*correspondent author

Abstract

Cancer is one of the major causes of non-accidental death in human. Early diagnosis of the disease allows clinician to administer suitable treatment, and can improve the patient's survival rate. Traditional diagnosis involves trained clinicians to visually examine the respective medical images for any signs of nodule development in the body. However due to the large scale of the medical image data, this manual diagnosis is often laborious and can be highly subjective due to inter-observer variability. Inspired by the advanced computing technology which is capable of performing complex image processing and machine learning, researches had been carried out in the past few decades to develop computer aided diagnosis (CAD) systems to assist clinicians detecting different forms of cancer. This paper reviews computer vision techniques adopted in medical image analysis, in particular, for cancer detection. The review focused on the detection of the most common form of cancer types, namely breast cancer, prostate cancer, lung cancer and skin cancer. A recent proposed cloud computing frame work has inspired the researchers to utilize the existing works on image based cancer research and develop a more versatile CAD system for detection.

Keywords: Segmentation; Computer vision; CAD system; Breast cancer, Lung cancer; Prostate cancer; Skin cancer; Cancer detection; Computed Tomography (CT); Ultrasound; Mammogram; Dermatography. Visual features,

1. Introduction

Cancer is a type of disease in which a group of cells exhibits irregular cell growth cycle. In the a normal cell cycle, the cells undergo mitosis process to replicate itself and hence the cell grows (Lee and Chen 2013a; Nahar et al 2011); eventually the *programmed cell death* process called apoptosis leads the cells to die in order to regulate its growth. In cancer, the cells lost such balance and grow uncontrollably, to form malignant tumors invading the surrounding tissues. The cancer cell can also migrate to other parts of the body by the bloodstream or lymphatic system, and continue to spread from the new location.

The cause of cancer has not been fully unveiled, however certain habits, such as smoking, exposure to radiations and environmental pollutants are known to cause cancer. Inherited genetic defects are also linked to the cause of some cancers (Nahar et al 2011; Rodrigues et al. 2006; Avila-Garcia et al. 2008; Doi 2005). Early diagnosis of the cancer will allow the clinician to remove the cancer cells via operation or administer

suitable treatment plan to eliminate the cancerous cells using chemical or radiation treatment.

Traditionally, cancer can be detected from the presence of certain symptoms, such as irregular markings on the skin, or hard lumps on the body. In clinics, screening tests (Chen and Chen 2006; Pisano et al. 2005; Joshua et al. 2013; Wu et al. 2011) and medical imaging (Nahar et al. 2012; Lee and Chen 2013b; Tang et al 2009; Huang et al 2007; Huang et al 2012) are the initial stages in cancer detection. Once a suspected cancer has been detected, tissue samples from the suspected region are extracted and examined (Nahar et al. 2012; Nahar 2007; Chen and Chen 2006; Hung et al 2011).

Non-invasive diagnosis of cancer involves a trained clinician visually examine different types of medical images, and identify the possible locations which resembles signs of malignant tumors . The accuracy of the diagnosis is highly dependent on the experiences of the clinician. Furthermore, with a large volume of medical database, this process is laborious and hardly consistent (Chen et al. 2007; Cheng et al 2010; Cruz-Roa, et al 2011; Cho et al 2010).

With the advance in digital computing technology, many researchers have combined image processing, pattern recognition, and artificial neural network to develop computer aided diagnosis (CAD) systems to assist the clinicians in the diagnosis process (Rolim et al 2010; Lee and Chen 2013b; Verma & Zako 2001; Ye et al. 2009) Figure 1 shows a general framework for a CAD system for skin cancer detection (Lee and Chen 2013b).

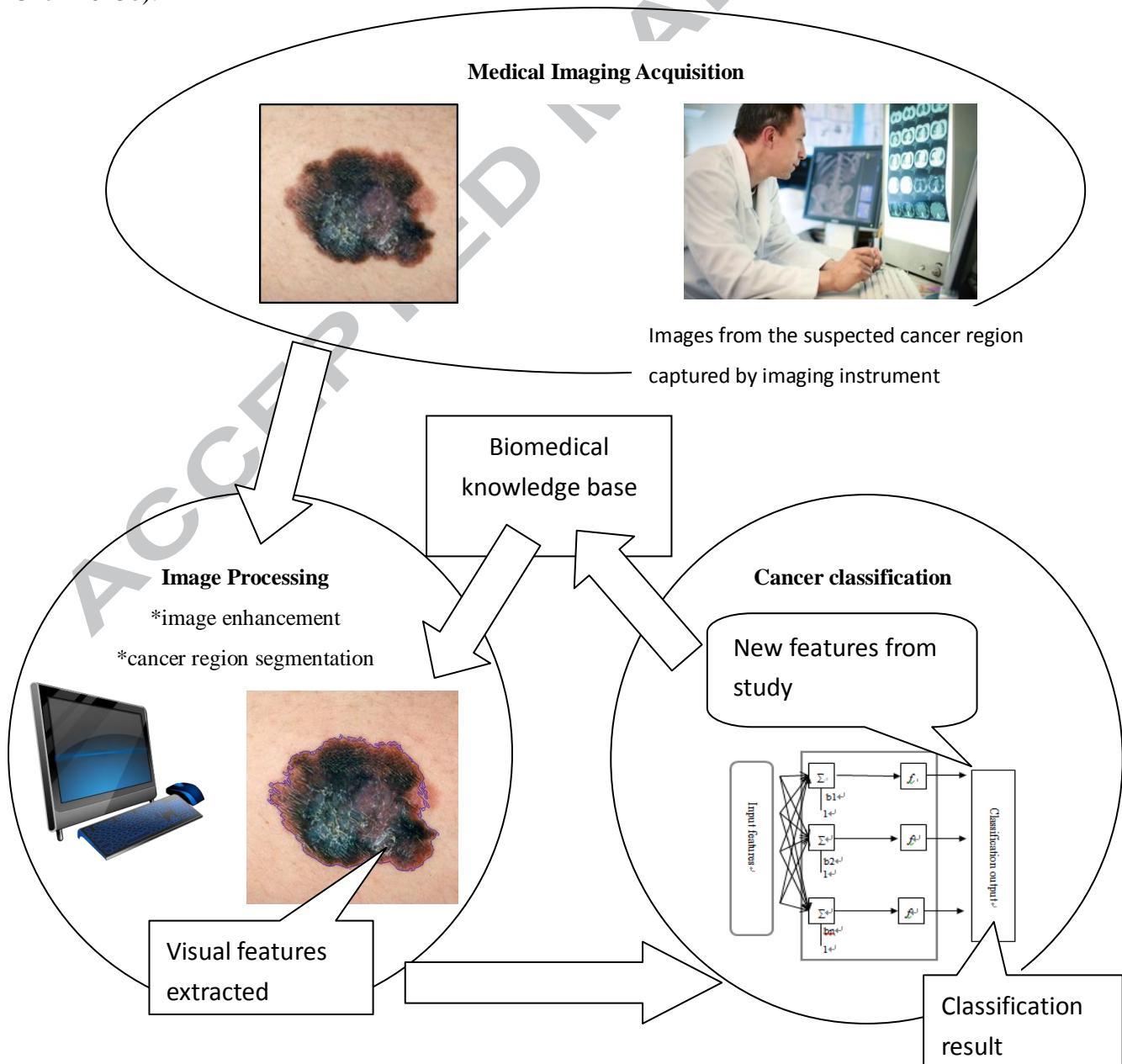


Figure 1 Framework for the cancer diagnosis system

In a CAD system for cancer detection, medical images are recorded using the appropriate imaging systems. The image acquisition device can be adjusted to ensure the consistent image quality under the same laboratory environment (Lee and Chen 2013a; Zheng et al. 2006). The captured images undergo series of software-based algorithms to isolate the suspected cancerous regions from the rest of the image (Lee and Chen 2013; Tong et al. 2010). A biomedical knowledge based features, such as textures (Rodrigues et al. 2006; Tang and Guo 2011), and shapes (Lee and Chen 2013a; Tourassi et al. 2006; Soysal & Chen 2008; Celebi et al. 2005) can be extracted to characterise the extracted segments. This feature space forms a biometric describing the suspected region. A supervised classifier can be implemented and trained using the existing sample images, to learn the distinctive patterns associated with the biometric. The diagnosis can be performed by comparing the feature patterns between the test sample and the trained patterns in the classifier features (Lee and Chen 2013; Joshua et al 2013; Sharaf-elDeen, et al. 2013).. Newly developed features from the study can be fed back to the biomedical knowledge database to improve the existing feature sets (Nahar et al 2007; Peng et al 2006).

Image segmentation plays a crucial role in a CAD system. It aims to isolate the suspected region from the rest of the image (Lee and Chen 2013b; Cheng et al 2013). It can also incorporate visual features such as color and texture information with other statistical and biological features to distinguish different regions in an image. Hence the unaffected regions can be removed and leaving the suspected regions which resembles certain visual patterns, such as irregular texture (Padmanabhan and Sundararajan 2012;), color and intensity (Umbaugh et al. 1989; Yuksel & Borlu 2009; Lee and Chen 2013b). An accurate segmentation result will help to determine the location and the size of the tumor, which is important for treatment planning.

The major challenge in this field of research is to build a fully automatic CAD system which can analyze large quantities of images to provide an accurate diagnosis and at the same time, robust enough to handle the biological variations in humans (Spurgeon 2005; Lee and Chen 2013a).

In this paper, we are going to provide an overview of the segmentation process used in these common medical image modalities to detect the most common forms of cancer: breast cancer, prostate cancer, lung cancer and skin cancer. We also investigate the recent work on these cancer detections with medical image processing and analysis. We discuss how image processing techniques have assist the clinicians in IMRT for cancer treatment. To utilize the existing research works on the algorithms to segment and classify for cancer images, a cloud computing frame work for image based cancer research has been proposed. In this paper, we discussed the advantages of this frame work and future works involved in developing a cloud computing based CAD system for cancer research.

2. Breast Cancer Detection

Breast cancer is a type of cancer originating from breast tissue, and it accounts for 23% of all cancers in women (Berman 2007). The most effective way to detect breast cancer is through the breast mammogram screening, however the major limitation for mammography diagnosis is sensitivity. This diagnosis is less sensitive in younger women especially cancers in dense breasts become difficult to detect (Smith et al. 2004;

Baker 2003).

Mammography is the most common imaging technique to detect breast cancer. (Boukerroui et al 2003; Chang et al. 2003; Chen et al. 2002; Cheng et al. 2010). In contrary to x-ray mammography, the conventional B-mode ultrasound is used to distinguish benign masses from malignant cancerous masses. Clinical studies have tried to characterize these masses by the texture and geometric properties (Arger et al. 2001; Huang et al. 2012; Jo et al. 2013). Figure 2 illustrates an example of breast cancer detection in mammography and ultrasound.

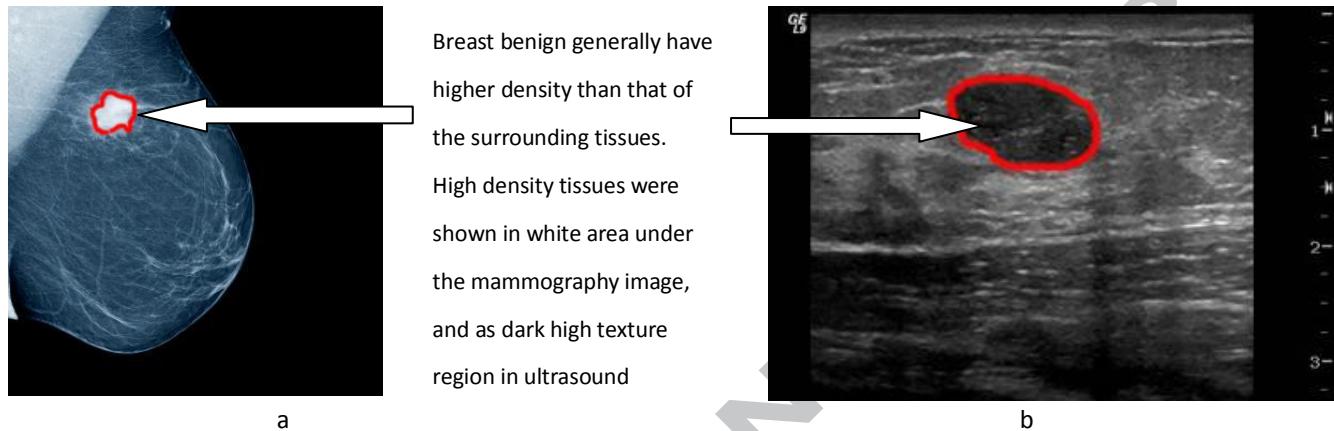


Figure 2: Breast cancer image from (a) mammography, and (b) ultrasound: showing the region where a benign has been detected in the breast (indicated with an arrow) – The suspected sighting of the breast benign has shown distinctive features compare to the surrounding normal tissues. (images are adopted from welcome images and Radiologyinfo.org)

A suspected benign can be detected by observing the regions showing abnormal high density. The texture and shape features are commonly used to detect the presence of breast cancer (Baker 2003), however one of the major hurdles in detecting breast cancer in mammography is that the high breast density, usually in young age group, may be difficult to detect the benign in the early stage, when the size is small. In conjunction with mammography, ultrasound technique has also been used to detect the presence of the benign in breast.

Recently, intensity-modulated radiation therapy (IMRT) has been applied for the treatment of breast cancer (Hoppe et al. 2012; Popescu et al 2010). Clinical studies have shown improvement on dose distribution throughout the breast, and minimize the acute skin reactions for the patients (Popescu et al 2010). However, this treatment has not been widely administered in some clinic institutes because the treatment involving CT scan to be performed for each patient, and target volume need to be delineated, all these are time consuming and laborious (Hurkmans et al. 2001). The treatment will involve will require daily expose to the radiation dose for the treatment period. This may increase the risk of damaging the surrounding tissues and developing secondary cancer. Despite the effectiveness in dose distribution, the survival rate for IMRT compare to non-IMRT techniques remains an open question (Hoppe et al 2012; Arger et al 2001).

2.1 X-ray mammograms

In the image-based CAD for breast cancer, there are two different goals in segmentation of micro-calcifications: to obtain the locations of the suspicious areas to assist clinician for diagnose; and to classify the abnormalities into benign or malignant, which is a cancerous mass (Baker 2003; Berman 2007;

Cheng et al. 2010). The major types of image segmentation process commonly used in the CAD for breast cancer are: *low level thresholding; region based techniques; Mathematical morphology and texture based segmentation.*

2.1.1 Low level thresholding

Low level thresholding technique utilized the low level image information such as local pixel values embedded in the mammogram images to segment images into different regions. This technique is based on the expected bimodal intensity distribution in the selected window which contains the sub-images to be segmented. (Baker 2003; Ayres & Rangayyan 2007; Berman 2007) This approach does not require prior information for the thresholding based on the image intensity. However the lack of spatial characteristics means it is not effective to images with high noise.

2.1.2 Region based approach:

Region growing is a common method of segmentation. It groups the near-by pixels with the properties similar to those of a seed pixel. Muralidhar et al. (2010) used a model based active contour approach to delineating breast cancer. Baker (2003) and Burhenne et al. (2000) described and compared local thresholding and region growing methods, and had shown that local thresholding has greater stability but is more dependent on parameter selection. However the region based process requires image with high contrast and require initial seeds. The process can be time consuming.

2.1.3 Mathematical morphology:

Edge detection is a traditional method for segmentation. Many operators such as Gabor operators, Sobel gradient, Prewitt gradient and Laplacian operator have been developed to determine closed edge in image (Zheng 2010), and hence delineate the image into regions. Other mathematical morphological operations such as erosion, top-hat transformation and complex morphological filters with multi-structure elements can also be used with adaptive neuro-fuzzy inference classification techniques (Huang et al. 2012). This method is generally efficient for analyzing geometric aspect of the image. However it generally requires a priori knowledge of the resolution level of mammograms in order to determine the sizes and shapes of the structure elements.

2.1.4 Texture-based segmentation

Image texture analysis has also been used in CAD system for feature enhancement, segmentation and classification. Texture features such as multi-resolution wavelet features (Chen et al. 2002; Mencattini et al. 2008), or Gabor features (Jo et. al. 2013, Zheng 2010) can be utilized to distinguish the micro-calcifications (small object) and normal tissues (large objects). The multi-resolution wavelet representation provides a natural hierarchy to embed an interactive paradigm for accomplishing scale-space feature analysis. The common scheme for wavelet transform on the detection of micro-calcification is to reconstruct the image from the transform coefficients modified at each level by local and global non-linear operators (Mencattini et al. 2008). Jo et al. (2013) has proposed a breast cancer detection algorithm based on the texture properties of the mass area. They measure the homogeneity using support vector machine to analyze the texture properties of the selected mass area. Zheng (2010) developed the Gabor Cancer Detection (GCD) algorithm, with utilized Gabor features. A set of edge histogram descriptor are extracted as texture descriptor for

analysis. To achieve maximum performance, it is important to study two crucial issues: wavelet base and non-linear functions of the wavelet coefficients.

2.2 Ultrasound imaging

Ultrasound is a non-invasive and mobile medical image modality to detect masses in the breast. Ultrasound imaging of the breast is typically done as an addition to x-ray mammography when breast cancer is suspected. These masses can be classified into benign masses and malignant masses by study their texture and geometric properties (Arger et al. 2001; Boukerroui et. al. 2007; Horsch et al. 2006). Inspired by these studies, researchers are seeking for an image-driven CAD system for automatic detection and characterization of breast cancer.

(Horsch et al 2004, 2006) has presented a method using threshold technique on the images with pre-enhanced mass structures. They determined and compared the effectiveness of various features using linear discriminant analysis, (Horsch et al. 2006; Drukker et al. 2013) and found the best two features, namely, depth-to-width ration (shape), and normalized radial gradient (margin), to distinguish benign and malignant masses.

Huang and Chen (2004) proposed a combination of neural network classification and watershed segmentation methods to extract the contours of the breast tumor in ultrasound image. In this approach, a self-organizing map (SOM) texture-based neural network is used to adaptively select appropriate preprocessing filters, which improves the effectiveness of the watershed segmentation algorithm. Madabhushi and Metaxas (2003) combined intensity, texture information, and empirical knowledge adopted by the clinicians with a deformable shape model and has automated the segmentation process, despite training is a prerequisite. Table 1 outlines the comparisons of these segmentation techniques for breast cancer detection, the image modalities used for these algorithms and related works.

2.3 Intensity-modulated radiation therapy (IMRT)

Radiotherapy is an effective treatment modality for early stage breast cancer. Despite the apparent success from multiple of clinical studies (Hoppe et al 2012), other studies also demonstrate significant dose inhomogeneity in the superior and inferior regions of the breast when applying wadged beam in the conventional treatment (Buchholz et al. 1997; Senkus-Konefka & Jessem 2007). Therefore it is important that an optimum solution for whole-breast radiotherapy should improve on dose homogeneity.

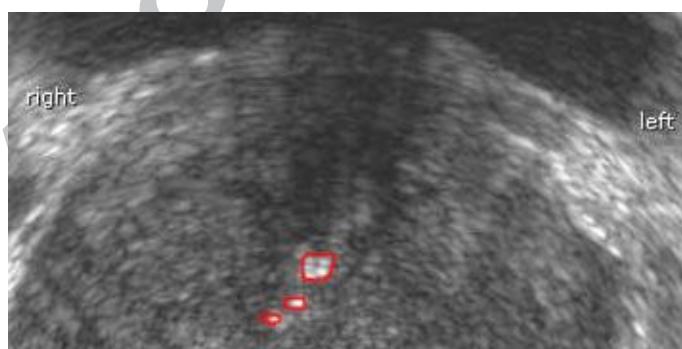
The role of IMRT is well established due to its tumor doses escalation characteristics. The goal is to deliver high dosage of radiation to the tumor while sparing the surrounding tissues to minimize damages to the normal tissues. Although there are limited literature describing image processing and segmentation used in this modality, especially for breast cancer treatment. It is evident that to perform an accurate diagnosis and treatment for IMRT, a precise model of the anatomical structure and the location of the tumor need to be constructed. Studies have shown the long term side effects of the IMRT treatment for breast cancer to the surrounding tissues and organs (Senkus-Konefka & Jessem 2007; Hoppe et al. 2012) Furthermore, due to the slow rotation of the IMRT system, motion artifacts have become apparent in the IMRT imaging, and may affect the precision of the treatment.

Segmentation techniques	algorithms	Image motility
Statistical method using thresholding: Use global low level features or local statistic information	Geometric constrained phase model (Ayres & Rangayyan 2007) Support vector machine for intensity features (Chang, 2003) Multi-resolution analysis (Nakayama et al. 2007)	mammography
Region based approach:	3D region grow technique (Horsch et al. 2004) Watershed algorithm (Huang & Chen 2004)	ultrasound
Mathematical morphology:	Model based active contour (<i>Muralidhar et al. 2010</i>) 3D discrete active contour (Chang, et al. 2003)	mammography; ultrasound
Texture-based segmentation:	Dynamic wavelet processing (Mencattini et al. 2008) Wavelet transform & neural network (Chen et al. 2002) Gabor operator for breast density pattern (Kinoshita, et al. 2007) Gabor filters (Zheng, 2010)	mammography ultrasound

Table 1: Comparison of segmentation methods for breast cancer detection

3. Prostate Cancer Detection

Prostate cancer is the most commonly diagnosed cancer in adult male populations. Early detection and intervention of progressive prostate cancer may help to improve the survival rate (Lim et al. 2008; Hricak et al. 2007). The key to the diagnosis, treatment and monitoring of this disease depends on the accurate segmentation of the prostate volumes and boundaries. Ultrasound has been the main imaging modality for prostate related applications because it is cost effective, innocuous and can be monitored in real time (Bloch et al. 2007; Campadelli, et al. 2010; Sung et al. 2011). However the major short coming of this image modality is its low signal-to-noise ratio and the presence of speckle noise make it difficult for automatic segmentation of the ultrasound images (Sung et al 2011; Vos et al. 2008; Hricak et al. 2007; Lim et al. 2008). Figure 3 depict the prostate cancer nodules in an ultrasound image.



The prostate nodules are generally small, and in ultrasound imaging, these nodules are difficult to detect due to high level of speckle noise present in ultrasound images.

Figure 3: Prostate cancer image from ultrasound: indicating in red, showing the region where small cancerous nodules have formed in the prostate region (image adopted from Radiologyinfo.org)

Classical techniques have been proposed for prostate segmentation. Sung et al. (2011) used the baseline and peak intensities of the image signal in conjunction with other perfusion parameters to determine the size of prostate cancer using a supported vector machine. Derivative edge detection techniques and non-linear filtering (minimum/maximum filtering) techniques were also introduced in (Heijmink et al. 2007), to obtain the second derivative and the gradient images which represent the possible edges of the prostate. However

these methods have not used shape modeling and have undergone limited validation.

Secondary features have also been used in detecting prostate boundary. Statistical moments have been taken into account for nature of TRUS image by considering the probe position in the shape normalization step and by making the image feature descriptor invariant to probe rotation (Choi et al. 2007; Doi 2005). Liu et al. (2009) proposed an unsupervised segmentation method, using fuzzy Markov random fields for the segmentation of multispectral MR prostate images. In their study, both hard and fuzzy MRF models have two groups of parameters to be estimated: the MRF parameters and class parameters for each pixel in the image. Manaco et al. (2010) used probabilistic pairwise Markov models to detect regions of carcinoma in the prostate region.

Recent research has focused on the including prior information about shape and speckle models, and delineate prostate in for 3D guided biopsy for prostate cancer. Fei et al. (2011) developed an automatic segmentation method based wavelet transform for 3D TRUS images of the prostate, and also developed a non-rigid registration algorithm for TRUS and PET/CT images. Yin et al. (2012), developed an automated CG segmentation algorithm based on Layered Optimal Graph Image Segmentation of Multiple Objects and Surfaces (LOGISMOS). 3D segmentation and registration on the prostate regions allows the CAD to automatically detect the boundary of the prostate region and accurately guide the biopsy system to the target region. Table 2 shows the recent work on prostate cancer detection algorithm and the corresponding prostate cancer image modalities

Detection techniques and algorithms	Prostate cancer image modalities
SVM with pixel intensity in conjunction with perfusion parameters (Sung, et al. 2011)	TRUS
Derivative edge detection and non-linear filtering (Heijmink et al. 2007)	TRUS
Fuzzy Markov random fields (Liu et al. 2009)	MRI
Probabilistic pairwise Markov models (Manaco et al. 2010)	MRI
Shape and speckle model using wavelet transform (Fei et al. 2011)	3D TRUS, TRUS and PET/CT
LOGISMOS for segmentation and registration (Yin et al. 2012)	3D MRI

Table 2: Recent works on prostate cancer detection algorithms.

4. Lung Cancer Detection

Computed tomography (CT) is the most sensitive and specific diagnostic modality for detecting lung nodules. The reliable detection of such nodules is crucial for early detection of lung cancer and other nodular lung diseases. (Chien et al. 2008; Oh, et al. 2011; Wu et al. 2011) Automatic detection of lung nodules is the most study problem in computer analysis of chest radiographs. Nodules show up as relatively low-contrast white circular objects within the lung fields, and are difficult to detect due to overlapping shadows from other structures such as vessels and ribs (Wu et al. 2011; Ye et al. 2009). Figure 4 illustrates an example of the lung cancer detection in X-ray and chest CT images.

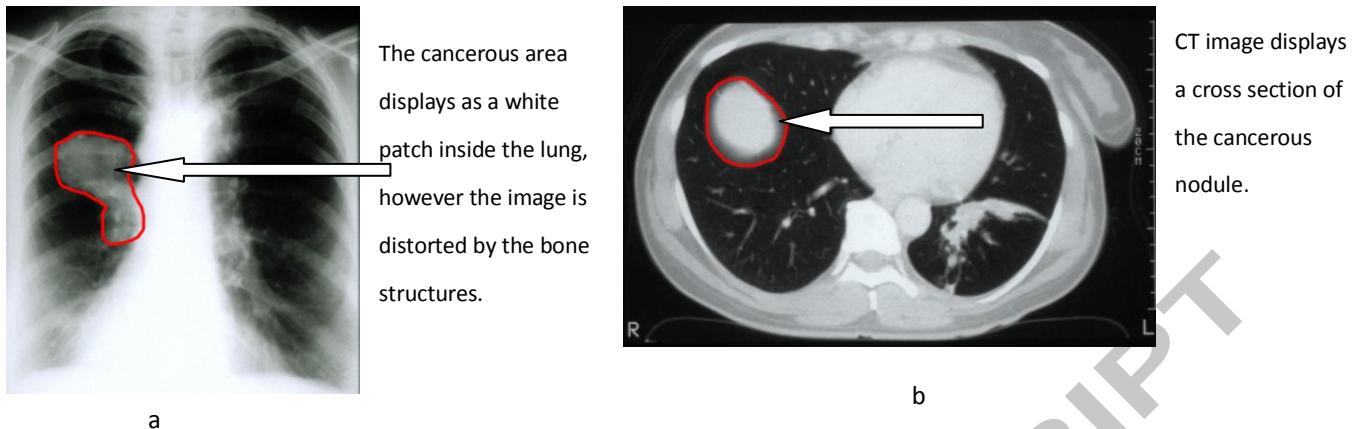


Figure 4: Lung cancer image from (a) X-ray, and (b) CT imaging: showing the region where cancer (indicated with an arrow) has been formed in lung. (images adopted from welcome images)

In chest x-ray image, the cancer region has been shown as a massive white region in the lung area. Typical feature used to detect lung cancer in X-ray is the image illumination, and the size of the effected region. However, the major issue facing in lung cancer segmentation in X-ray imaging is that the overlapping of the ribcage makes the lung nodules, an early form of lung cancer, difficult to detect (Tanino, et al. 2003) Chest CT images, on the other hand, displays the horizontal slides of the lung to detect the presence of the lung nodules developed in the lung (figure 4b). The CT image offers the clear image quality without overlapping distortion as shown in the chest X-ray. However the presence of air packets in the lungs can cause errors in lung nodule detection (Li, 2007). Automatic detection of lung nodules consists of two steps: 1) initial candidate nodules detection and 2) false positive reduction (Li, 2007).

4.1 Nodule candidate detection

Several methods applied filters to enhance the nodules or eliminate background by blurring (Zhao et al. 2004; Linguraru et al. 2006). Nodule candidates are detected using template marching or a modified Hough transform (Tanino et al. 2003; Ye et al 2009). In other cases the nodule is detected by thresholding techniques (Ko and Betke 2001; Zhao et al. 2004). These methods utilise the global intensity information to identify the lung nodules, however they requires images with high contrast and are vulnerable to image distortions. Boyce et al. (2013) investigated the lung nodule detection with the use of stereoscopic visualization compared with the standard posero-anterior images. Their preliminary result has shown that the use of stereoscopic visualization reduced the sensitivity for nodule detection, however it improved the positive predictive value.

4.2 Reduction of false positives

A related problem is to classify nodules into benign or malignant. Supervised learning scheme has been commonly used in classification and reduces FPs. Clinically identified samples have been used to determine benign and malignant nodules by evaluating the correlation between the test images and nodule samples (Oh et al 2011; Samei et al. 2007; Ye, et al. 2009; Ge et al. 2010). They had performed an impressive result but such CAD systems require massive training and learning. Classifiers such as nearest neighbor (Tanino et al. 2003; Zhao et al. 2004) require human interaction to initiate the clusters. Table 3 outlines the common classification scheme used for classifying lung nodules.

Classification scheme	Algorithms used	Image modalities
Supervised Learning	Bayesian network (Oh, et al. 2011) 3D gradient field and 3D ellipsoid fitting (Ge, et al. 2010) SVM (Ye, et al 2009)	X-ray CT 3D-CT,
Clustering	VNQ filters and PCS-based classification (Tanino et al. 2003) LDM algorithm (Zhao, et al. 2004)	3CT

Table 3: Common classification schemes for lung nodule classification

4.3 IMRT in lung cancer treatment

The application of IMRT in lung cancer treatment is a relatively new procedure. The effectiveness of this procedure lies heavily on the accurate target volume delineation. Inaccurate target delineation can compromise the treatment outcome and may also damage the surrounding normal lung tissue, and other critical organs. An approach based on FDG-PET techniques for guided IMRT had been introduced by Song et al. (2006). This technique use both CT and PET/CT images to construct the cross section of the lung and isolate the tumor location. The preliminary result shows improvement on target volume delineation.

5 Skin Cancer Detection

Skin cancers can be classified into melanoma and non-melanoma. Although melanomas are much less common, they account for most of the mortality from skin cancer. Detection of malignant melanoma in its early stages considerably reduces morbidity and mortality (Carrara et al. 2007; Lee & Chen 2013a). Dermatoscopy is a non-invasive imaging technique based on oil immersion, which renders the skin translucent, hence allowing, and therefore offers clear visualization of surface and subsurface structures. However to automatic segment a skin image into lesions is a challenge task. Figure 5 presents typical dermatoscopy images for skin cancer detection.

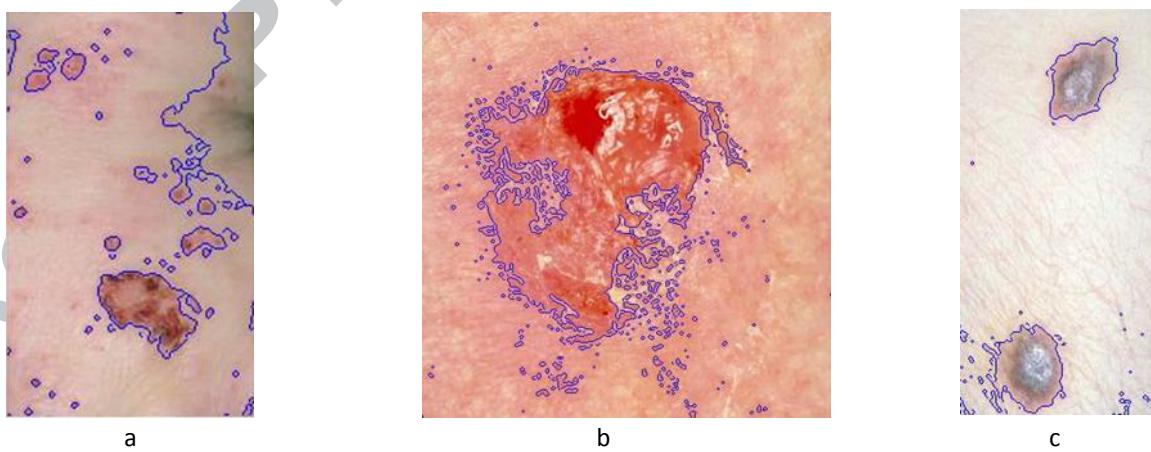


Figure 5: Skin dermatoscopy images for (a) Squamous cell carcinoma; (b) Basal cell carcinoma and (c) Malignant melanoma. These three types of skin cancer display different visual characteristics.

The squamous cell carcinoma (Figure 5a) and basal cell carcinoma (Figure 5b) are the most common forms of skin cancer. The squamous cell carcinoma is commonly a red, thickened patch at the sun-exposed skin, this skin cancer type usually effecting a larger area of the skin. Basal cell carcinoma, on the other hand, usually presents as a raised, smooth bump on the skin. However the most lethal form of skin cancer is

malignant melanoma, (Figure 5c) it appears in dark colour and usually in the irregular shapes (Lee and Chen 2013b) .

5.1 Skin cancer segmentation

The changes in the pigment of the skin have been the key features to detect skin cancer. However the quality of the skin dermatography is highly influenced by the environment, such as lighting and instrument. Uneven lightings may cause the shadows and uneven illumination on the skin, and cause error in segmentation. Furthermore, segmentation with based only on the color features may not be robust to address the variations in nature skin colors. Other features such as asymmetry, border irregularity, and area are also used in clinical diagnosis. Researches have been conducted to combine these clinical features to detect skin cancer automatically (Carrara et al. 2007).

5.2 Recent works in automatic skin cancer image extraction

Celebi et al. (2005) combined color quantization and region growing method to segment skin lesions under the unsupervised learning algorithm to detect the border of the infected skin cancer. Carrara et al. (2007) investigated the use texture information embedded in the image with a supervised learning algorithm to assist clinicians to detect skin cancer with varying skin pigments. These approaches are less dependent on the color information of the dermatoscopy images, and can address the limitation of the color clustering based approaches, such as the similar color tones of the hair and the lesions.. Despite the color features embedded in the skin images, edge detection algorithm which uses the texture variations between the normal and cancerous skin regions has also been investigated (Tang and Guo, 2011). Lee and Chen (2013b) used the color constancy approach to address the color tone variations due to different skin samples and image capturing environment. Using the type-2 fuzzy set algorithm to determine an optimum threshold level, this approach also allow to segment three common types of skin cancer, namely basal cell carcinoma, squamous cell carcinoma and malignant melanoma. Comparisons of the recent works on skin cancer segmentation has been summarized in Table 4.

Recent works	Features used	Algorithm	cancer detection
Celebi et al. 2005	HSV color,	Color quantization Region growing	Malignant melanoma
Carrara et al. 2007	RGB color, texture	Supervised ANN Gabor filters	Malignant melanoma Squamous cell carcinoma
Tang & Guo, 2011	Color, texture	Wavelet diffusion, active contour	Malignant melanoma
Lee & Chen, 2013b	RGB color	Optimum threshold using type-2 Fuzzy set	Basel cell carcinoma Squamous cell carcinoma Malignant melanoma

Table 4: Recent work in skin cancer segmentation algorithms

6 Future research directions in CAD for cancer detection

The image processing based CAD for cancer diagnosis has only been made to detect singular cancer type. Research in image processing for cancer segmentation had utilized the low level features embedded in the image to detect the presence of cancer cell. Depending on the image modality used, multiple visual features usually used in conjunction with some optimum classification algorithm to obtain the best segmentation result.

6.1 Image features in cancer segmentation

As discussed in the previous sections, cancer images from different modalities had provided different image characteristics to distinguish cancerous regions from the normal tissue. In automatic cancer image segmentation, low level image features have been widely used to represent these visual differences in the cancerous region.

In X-ray imaging modalities, such as mammography and chest X-ray, the density of the body tissue is shown in high intensity. Detection on the cancerous cells in breast and lung in X-ray can be influenced by the density of the body tissue. In ultrasound imaging technique, the texture and density of the body tissue can be detected by the resonance of the sound wave. These images generally display the shape of the organs or tumours. However ultrasound generally have poor image quality, precise boundary of the cancerous region always difficult to obtain, and it is difficult to identify small nodules. Computer tomography (CT) takes series of images showing horizontal slides of the region. The image quality displays the shape and density of the organ tissues. Image intensities and shapes are the common features used to detect cancerous region in this modality. Skin cancer images are obtained by dermatography, which displays the full colour information about the skin pigment. Cancerous regions are detected based on the colour variations. Table 5 summarised the features used in different image modalities and the targeting cancer type detection

Image modality	Image features	Cancer types	Major issues
X-ray	intensity, texture	breast; lung	highly affected by tissue density and overlapping organs and bone structure
Ultrasound	texture, shape	breast; prostate	difficult to detect small nodules and accurate border of the tumour
Computer tomography	intensity, shape; texture	lung; prostate	sensitive to other structure in the organ, such as blood vessels and air packets in lungs.
Dermatography	colour, shape	skin	highly affected by the skin tone variations and ambient light source.

Table 5: Comparison of cancer imaging modalities for different cancer detection.

In addition, semantic features such as texture pattern definition, the size of the region, and shape definition can provide extra biological characteristics in segmentation. These semantic features have been used to identify or track particular object in multimedia data (Tjondronegoro & Chen 2010), and can also provide key features in cancer segmentation.

Multiple modalities analysis have also been deployed in medical practice to detect cancer (Horsch et al 2006). However in the CAD for cancer detection, it only utilise single diagnosis from single modality, and

sometimes the result may be contradictory. Extension for image processing in cancer detection, can combine critical features extracted from images in different modality to improve the detection of cancer in terms of its location, size and shape.

6.2 Cloud computing in CAD design for cancer detection

A cloud computing framework design for cancer research has been proposed based on the Microsoft cloud service technology, (Avila-Garcia, et. al. 2008) to utilize the massive existing algorithms developed by different researchers. Figure 6 shows the proposed cloud computing frame work for image-based cancer detection.

In cloud computing, the image database and the existing image processing algorithms are made available upon the user queries (Rolim et. al. 2010, Vecchiola et. al. 2009). This allows the clinicians to access the most suitable algorithm to process the data image according to different cancer type and different image modalities that were captured. Clinicians can access and update the new findings in cancer research, such as new features to detect cancer. The access to different image database allows the clinicians to consult with different case studies or collaborate with other researchers of the same field. In technical aspect, this frame work saves the time and resources of implementing pre-existing algorithms for analyzing different image data. New image data captured using new devices can be made available for all researchers and make the development of new algorithms more efficient.

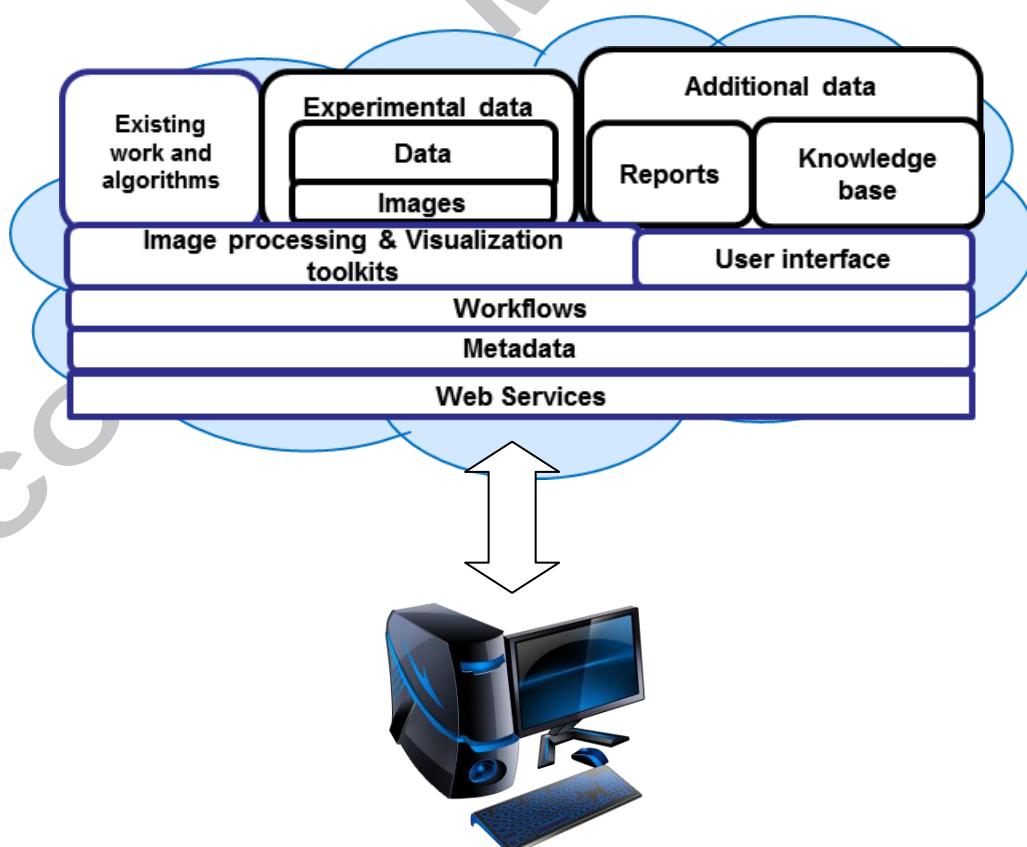


Figure 6: Cloud computing framework for image-based cancer detection. (adopted from Avila-Garcia, et. al. 2008)

The challenge in developing such framework include building a repository for the existing algorithms; developing an effective metadata information for cancer images; define the work flow for accessing sensitive medical data (Avila-Gacia et al. 2008, He et al 2010).

7 Conclusion

In this survey of image based CAD system for cancer detection, we had studied the major image modalities for detecting different cancer types in the medical field. We compared different segmentation algorithms for delineating different cancerous region, and outlined the advantages and disadvantages in each segmentation algorithms. To correctly identifying the cancer cell, relevant features are extracted and studied. In this paper, we examined the major visual features used to detect the four major types of cancer, namely, breast cancer, lung cancer, prostate cancer and skin cancer. In this paper, we also discussed and compared the recent studies in image based cancer segmentation and detection algorithms. The visual features comparison for different cancer types and the corresponding medical image modalities for cancer detection allow the researchers to compare the characteristics of different cancers and the existing image processing approaches in cancer detection.

Image segmentation had proved to be very effective in detecting cancerous tumors in different medical imaging modalities, however due to variations of biological information on different parts of the human anatomy, human intervention is almost inevitable. It requires experienced clinician to either provide training data, or set up the initial conditions for classification. Researches have been done to detect different types of cancer based on the visual information extracted from medical images. However most of the image based cancer diagnosis systems only detect singular type of cancer, and most image processing algorithms are customized to address particular set of image data, or images from the same modality. Recently a cloud computing framework has been proposed to incorporate algorithms developed by different researchers for cancer image detection. The framework will encourage the researcher to collaborate with clinicians to develop a versatile CAD system for cancer research. This cloud computing frame work can be served as the repositories for the existing image processing algorithms and cancer image databases. Providing a research and development platform for image based cancer research. This extension allows the future CAD design to be more versatile, and the newly discovered information can be shared more rapidly. The future extension in cloud based CAD design requires researchers to build a biomedical visual feature archive and algorithm archive that allows the cloud based CAD to utilize the existing algorithms and feature sets for cancer detection, and allows the newly developed algorithms, features and biological information to be updated. Another future extension for cancer detection is to use multiple image modalities for cancer study. Different types of image modalities offer different insights of the cancer region. A combination of the feature analysis from these features can improve the effectiveness and accuracy for cancer detection.

Another future research direction for image based cancer detection can incorporate biometrics of the individual patient, allows a customized medical diagnosis and treatment for cancer. Biometric information such as the size and thickness of the organs and tissues can help identify the exact location of the internal cancerous cells such as lung and breast carcinoma, and allows proper radiation dosage delivered to the cancerous site and minimizes the damages to the surrounding normal tissues. Natural skin tone of the individual patient can be used to accurately delineate the boundary of skin cancer. Allows the clinicians to determine the types of skin cancer, and monitor its propagation on the skin surface.

8 References

- Arger, P. H., Sehgal, C., Conant, E., Zuckerman, J., Rowling, S. E. & Paton, J. A. (2001) Inter-reader variability and predictive value of us descriptions of solid breast masses. *Acad. Radiol.*, 335-342.
- Avila-Garcia, M.S., Trefethen, A.E., Brady, M., Gleeson, F. and Goodman, D. (2008) Lowering the barrier to cancer imaging, *IEEE 4th International Conference on eScience*, 63-70.
- Ayres, FJ., and Rangayyan, RM., (2007) Reduction of false positives in the detection of architectural distortion in mammograms by using a geometrically constrained phase portrait model, *International Journal of Computer Assisted Radiology Surgery*. 1(6): 361-369.
- Baker, J.A., (2003) Computer-aided detection (CAD) in screening mammography: sensitivity of commercial CAD systems for detecting architectural distortion, *American Journal of Roentgenology*, 181:1083-1088.
- Berman, C.G., (2007) Recent advances in breast-specific imaging, *Cancer Control* 14(4):338-349
- Bloch BN., Furman-Haran, E., and Helbich, TH. (2007) Prostate cancer: accurate determination of extracapsular extension with high-spatial-resolution dynamic contrast-enhanced and T2-weighted MR imaging—initial results. *Radiology*, 245:176–185
- Boukerroui, D., Baskurt, A., Noble, J. A. & Basset, O. (2003) Segmentation of ultrasound images - multiresolution 2D and 3D algorithm based on global and local statistics. *Pattern Recognition Letters*, 24(4-5):779-790.
- Boyce, S., McAdams, H., Ravin, CE., Patz Jr., EF., Washinton, L., Martinez, S., Kowek, L., and Samei, E., (2013) Preliminary Evaluation of Biplane Correlation (BCI) Stereographic Imaging for Lung Nodule Detection, *Journal of Digital Imaging*, 26(1):109-114
- Buckley, DL., Roberts, C., Parker, GJ., Logue, JP., and Hutchinson, CE. (2004) Prostate cancer: evaluation of vascular characteristics with dynamic contrastenhanced T1-weighted MR imaging—initial experience. *Radiology* 233:709–715
- Burhenne, L. W., Wood, S. A., D'Orsi, C. J., Feig, S. A., Kopans, D. B., O'Shaughnessy, K. F.,(2000) Potential contribution of computer-aided detection to the sensitivity of screening mammography, *Radiology*, 215:554-62, 2000.
- Campadelli, P., Casiraghi, E., and Oratissoli, S., (2010) A Segmentation framework for abdominal organs from CT scans, *Artificial Intelligence in Medicine*, 50(1):3-11.
- Carrara, M., Bono A. and Bartolic., (2007) Multispectral imaging and artificial neural network: mimicking the management decision of the clinician facing pigmented skin lesion, *Physics, Medical Biology*, 149:2599-2613.
- Celebi, M. E., Aslandogan, Y. A. & Bergstresser, P. R. (2005) Unsupervised border detection of skin lesion images., *Proceedings of the International Conference on Information technology: Code and Computing*, 2:123-128.
- Chang, R. F., Wu, W. J., Moon, W., Chen, W. M., Lee, W. E. I. & Chen, D. R. (2003) Segmentation of breast tumor in three-dimensional ultrasound images using three-dimensional discrete active contour model. *Ultrasound in medicine & biology*, 29(11):1571-1581.
- Chang, R. F., Wu, W. J., Moon, W. K., Chou, Y. H. & Chen, D. R. (2003) Support vector machines for diagnosis of breast tumors on us images. *Academic Radiology*, 10(2):189-197.
- Cheikh, AB., Girouin, N., and Colombel, M., (2008) Evaluation of T2-weighted and dynamic contrast-enhanced MRI in localizing prostate cancer before repeat biopsy. *European Radiology*, 19:770–778

- Chen, C. H. & Lee, G. G. (1997) On digital mammogram segmentation and microcalcification detection using multiresolution wavelet analysis. *Graphical Models and Image Processing*, 59(5): 349-364.
- Chen, D. R., Chang, R. F., Kuo, W. J., Chen, M. C. & Huang, Y. L. (2002) Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks. *Ultrasound Medical Biology*, 28(10), 1301-1310.
- Chen Q. Chen YPP (2006), Mining Frequent Patterns for AMP-activated Protein Kinase Regulation on Skeletal Muscle, *BMC Bioinformatics*, 7:394-408
- Cheng SC, Cheng KY, Chen YPP (2013) GHT-Based Associative Memory Learning and Its Application to Human Action Detection and Classification, *Pattern Recognition*, 46(11): 3117–3128.
- Chien C-R, Chen TH-H. (2008) Mean sojourn time and effectiveness of mortality reduction for lung cancer screening with computed tomography. *International Journal of Cancer*. 122(11):2594-2599
- Choi YJ., Kim, JK., Kim, N., Kim, KW., Choi, EK., and Cho KS (2007) Functional MR imaging of prostate cancer, *RADiOGraphics*, 27:63-75
- Ciatto, S., Houssanmi, N., Gur D., Nishikawa, R., Schmidt, R., Metz, C., Ruiz, J., Feig, S., Birdwell, R., Linver, M., Fenton, J., Barlow, W., and Elmore, J. (2007) Computer-aided screening mammography, *New England Journal of Medicine*, 357(1): 83-85.
- Cruz-Roa, A., Caicedo, J.C., and Gonzalez, F.A. (2011) Visual pattern mining in histology image collections using bag of features, *Artificial Intelligence in Medicine*, 52(2):91-106
- Ding, M. & Fenster, A. (2003) A real-time biopsy needle segmentation technique using Hough transform. *Medical Physics*, 30: 2222-2233.
- Doi, K., (2005) Current status and future potential of computer-aided diagnosis in medical imaging, *The British Journal of Radiology*, 78: 3-19.
- Drukker, K., Horsch, K., Pesce, L., and Giger, M., (2013) Interreader Scoring Variability in an Observer Study Using Dual-Modality Imaging for Breast Cancer Detection in Women with Dense Breasts, *Academic Radiology*, 20(7):847-853
- Fei, B., Master, V., Nieh, P., Akbari, H., Yang, X., Fenster, A., and Schuster, D. (2011) A PET/CT Directed, 3D Ultrasound-Guided Biopsy System for Prostate Cancer, *Prostate Cancer Imaging – image Analysis and Image0Guided Interventions Lecture Notes in Computer Science*, 6963:100-108
- Feng, M., Moran, JM., Koelling, T., Chughtai, A., Chan, J., Freedman, L., Hayman, JA., Jaqsi, R., Jolly, S., Larouere, J., Soriano, J., Marsh, R., and Pierce, L., (2011) Development and validation of a heart atlas to study cardiac exposue to radiation following treatment for breast cancer, *International Journal of Radiation Oncology*, 79(1):10-18.
- Gavrielides, M. A., Lo, J. Y. & Floyd Jr, C. E. (2002) Parameter optimization of a computer-aided diagnosis scheme for the segmentation of micro-calcification clusters in mammograms. *Medical Physics*, 29:475-483.
- Ge, ZY., Sahiner, B., Chan, HP., Hadjiiski, LM., Cascade. PN., Bogot, N., Kazerooni, EA., and Zhou, C., (2005) Computer-aided detection of lung cancer nodules: False positive reduction using a 3D gradient field method and 3D ellipsoid fitting. *Medical Physics*, 32:2443-2454.
- He, C., Jin, X., Zhao, Z., and Xiang, T., (2010) A cloud computing solution for hospital information system, *Proceedings on IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS)* 2:517-520
- Heijmink SW., Futterer JJ., and Hambrock T., (2007) Prostate cancer: body-array versus endorectal coil MR imaging at 3T—comparison of image quality, localization, and staging performance. *Radiology*

- Hoppe, BS., Flampouri, S., Su, Z., Latif, N., Deng, NH., Lynch, J., Joyce, M., Sandler, E., Li, Z., and Mendenhall, N. (2012) Effective dose reduction to cardiac structures using protons compared with 3DCRT and IMRT in mediastinal Hodgkin lymphoma, *International Journal of Radiation Oncology*, 84(2):449-455.
- Horsch, K., Giger, M. L., Vyborny, C. J. & Venta, L. A. (2004) Performance of computer-aided diagnosis in the interpretation of lesions on breast sonography. *Academic Radiology*, 11(3), 272-280.
- Horsch K., Ginger, ML., Vyborny CJ., Lan L., Mendelson E. and Hendrick RE., (2006) Classification of breast lesions with multimodality computer-aided diagnosis: observer study result on an independent clinical data set, *Radiology*, 240(2):357-368
- Hricak H, Choyke PL, Eberhardt SC, Leibel SA, and Scardino PT. (2007) Imaging prostate cancer: a multidisciplinary perspective. *Radiology*, 243:28–53
- Huang, M.J., Chen, M.Y., and Lee, S.C. (2007) Integrating data mining with case-based reasoning for chronic diseases prognosis. *Expert Systems with Applications*, 32(3):856-867.
- Huang, M.L., Hung, Y.H., Lee, W.M., Li, R.K., and Wang T.H., (2012) Usage of case-based reasoning, neural network and adaptive neuro-fuzzy inference system classification techniques in breast cancer dataset classification diagnosis. *Journal of Medical Systems*, 36(2): 407-414.
- Huang, Y. L. & Chen, D. R. (2004) Watershed segmentation for breast tumor in 2-d sonography. *Ultrasound in Medicine & Biology*, 30(5):625-632.
- Hung, W.-L., Chen, D.-H., and Yang, M.-S., (2011) Suppressed fuzzy-soft learning vector quantization for MRI segmentation, *Artificial Intelligence in Medicine*, 52(1): 33-43.
- Kinoshita, SK., de Azevedo-Marques, PM., Pereira, RR., Rodrigues, JA., and Rangayyan, RM., (2007) Content-based retrieval of mammograms using visual features related to breast density patterns, *Journal of Digital Imaging*, 20(2): 172-190
- Lathan, CS, Neville, BA, Earle, CC. (2006) The effect of race on invasive staging and surgery in non-small-cell lung cancer. *Journal of Clinical Oncology*. 24(3):413-418
- Lee, H. and Chen, YPP, (2014) Cell cycle phase detection with cell deformation analysis, *Expert Systems with Applications*, 41(6): 2644–2651
- Lee, H. and Chen, YPP, (2014) Skin cancer extraction with optimum fuzzy thresholding technique, *Applied Intelligence*, 40(3): 415-426
- Lee, Y., Hara, T., Fujita, H., Itoh, S. & Ishigaki, T. (2001) Automated detection of pulmonary nodules in helical ct images basedon an improved template-matching technique. *IEEE Transactions on Medical Imaging*, 20(7): 595-604.
- Li, Q., and Doi, K., (2006) Analysis and minization of overtraining effect in rule-based classifiers for computer-aided diagnosis," *Medical Physics*, 33:320-328.
- Li Q. (2007) Recent progress in computer aided diagnosis of lung nodules on thin-section CT, *Computerized Medical Imaging and Graphics*, 31: 248-257
- Lim HK, Kim JK, Kim KA, Cho K-S. (2008) Prostate cancer: apparent diffusion coefficient map with T2-weighted images for detection—a multireader study. *Radiology*, 250:145–151
- Linguraru, MG., Marias, K., English, R., and Brady, M., (2006) A biologically inspired algorithm for microcalcification cluster detection, *Medical Image Analysis*, 10(6):850-862.
- Liu X., Langer, DL., Haider, MA., Yang, Y., Wernick, MN., and Yetik, IS. (2009) Prostate cancer segmentation with simultaneous estimation of Markov random field parameters and class, *IEEE Transaction*

on Medical Imaging, 28(6):906-915.

Madabhushi, A. & Metaxas, D. N. (2003) Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions. *IEEE Transactions on Medical Imaging*, 22(2):155-169.

McCann, J., Artinian V., Duhaime, L., Lewis, J., Kvale, P., and DiGiovine, B. (2005) Evaluation of the causes for racial disparity in surgical treatment of early stage lung cancer. *Chest Journal*. 128(5):3440-3446

Mencattini, A., Salmetr, M., Lojacono, R., Frigerio, M., and Caselli, F. (2008) Mammographic Images Enhancement and Denoising for Breast Cancer Detection Using Dyadic Wavelet Processing, *IEEE Transaction on Instrumentation and Measurement*, 57(7):1422-1430

Monaco, J., Tomaszewski, JE. Feldman, MD., Hagemann, I., Moradi, M., Mousavi, P., Boag, A., Davidson, C., and Madabhushi, A., (2010) High-throughput detection of prostate cancer in histological sections using probabilistic pairwise Markov models, *Medical Image Analysis*, 14(4):617-629.

Muralidhar, G.S., Bovik, A.C., Giese, J.D., Sampat, M.P., Whiteman, G.J., Haygood, T.M., Stephens, T.W. and Markey, M.K. (2010), Snakules: A Model-based Active Contour Algorithm for the Annotation of Spicules on Mammography, *IEEE Transactions on Medical Imaging*, 29(10): 1768-80.

Nahar, J., Imam, T., Tickle, K., Ali, S., and Chen, YPP. (2012) Computational Intelligence for Microarray Data and Biomedical Image Analysis for the Early Diagnosis of Breast Cancer, *Expert Systems With Applications*, 39 :12371–12377.

Nahar J., Tickle, KS, Ali S. Chen YPP (2011) Significant Cancer Prevention Factor Extraction: An Association Rule Discovery Approach, *Journal of Medical Systems* 35(3): 353-367, 2011.

Nahar J., Chen YPP, Ali S. (2007) Kernel Based Naive Bayes Classifier for Breast Cancer Prediction, *Journal of Biological Systems*, 15(1): 17-25.

Nakayama, R., Watanabe, R., Kawamura, T., Takada, T., Yamamoto, K., and Takeda, K. (2008) Computer-aided diagnosis scheme for the detection of architectural distortion on mammograms using multiresolution analysis. *International Journal of Computer Assisted Radiology Surgery*, 3(1):418-419.

Nishikawa, RM., (2007) Current status and future directions of computer-aided diagnosis in mammography, *Computerized Medical Imaging and Graphics*, 31:224-235.

Oh, JH., Craft, J., Lozi, RA., Vaidya, M., Meng, Y., Deasy, J., and Naqa, IE. (2011) A Bayesian network approach for modelling local failure in lung cancer, *Physics in Medicine and Biology*, 56(6):1635-1651

Pinsky, PF., (2004) An early- and late-stage convolution model for disease natural history. *Biometrics*, 60(1): 191-198.

Popescu, CC., Olivotto, IA., Bechham, WA., Ansbacher, W., Zavqorodni, S., Shaffer, R., Wai, ES., and Otto, K. (2010) Volumetric Modulated Arc Therapy Improves Dosimetry and Reduces Treatment Time Compared to Conventional Intensity-Modulated Radiotherapy for Locoregional Radiotherapy of Left-Sided Breast Cancer and Internal Mammary Nodes, *International Journal of Radiation Oncology*, 76(1):287-295.

Rangayyan, RM., and Ayres, FJ. (2006) Gabor filters and phase portraits for the detection of architectural distortion in mammograms, *Medical and Biological Engineering and Computing*, 44(10): 883-894

Rangayyan, RM., Prajna, S., Ayres, FJ., and Desautels, JEL., (2008) Detection of architectural distortion in mammograms acquired prior to the detection of breast cancer using Gabor filters, phase portraits, fractal dimension, and texture analysis, *International Journal on Computer Assisted Radiology Surgery*, 2(6): 347-361.

ACCEPTED MANUSCRIPT

- Rolim, C.O., Koch, F.L., Westphall, C.B., Werner, J., Fracalossi, A. and Salvador, G.S. (2010) A Cloud Computing Solution for Patient's Data Collection in Health Care Institutions, *Proceedings on the 2nd International Conference on eHealth, Telemedicine and Social Medicine*, 95-99
- Sahiner, B., Chan, H. P., Roubidoux, M. A., Helvie, M. A., Hadjiiski, L. M., Ramachandran, A., Paramagul, C., Lecarpentier, G. L., Nees, A. & Blane, C. (2004) Computerized characterization of breast masses on three-dimensional ultrasound volumes. *Medical Physics*, 31:744-754.
- Samei E., Stebbins, SA., Dobbins, JT., and Lo, JY. (2007) Multi-projection correlation imaging for improved detection of pulmonary nodules, *American Journal of Roentgenology*, 188(5):1239-1245
- Sampat, M. P., Bovik, A.C., Whitman, G.J., and Markey, M.K. (2008) A modelbased framework for the detection of spiculated masses on mammography, *Medical Physics*, 35:2110-2123.
- Schmid, P. (1999) Segmentation of digitized dermatoscopic images by two-dimensionalcolor clustering. *IEEE Transactions on Medical Imaging*, 18(2):164-171.
- Senkus-Konefka, E., and Jassem, J., (2007) Cardiovascular effect of breast cancer radiotherapy, *Cancer Treatment Review*, 23(6):578-593.
- Shen, L., Rangayyan, R. M. & Desautels, J. E. L. (1994) Detection and classification of mammographic calcifications. *State of the Art in Digital Mammographic Image Analysis*, 198-212.
- Smith, R. A., Duffy, S. W., Gabe, R., Tabar, L., Yen, A. M. F. & Chen, T. H. H. (2004) The randomized trials of breast cancer screening: What have we learned? *Radiologic Clinics of North America*, 42(5):793-806.
- Song, Y., Chan, M., Burman, C. & Cann, D. (2006) Inter-modality variation in gross tumor volume delineation in 18fdg-pet guided imrt treatment planning for lung cancer. *Proceedings on IEEE Conference on Engineering in Medicine and Biology*, 3803-3806
- Soysal, ÖM. and Chen, J. (2008) A new spectral feature for shape comparison, *Proceedings of the International Conference in Image Processing, Computer Vision, and Pattern Recognition*, 23-27.
- Sung, YS., Kwon, H-J., Park B-W., Cho, G., Lee, CK., Cho, K-S., and Kim, JK., (2011) Prostate cancer detection on dynamic contrasted-enhanced MRI: computer-aided diagnosis versus single perfusion parameter maps, *American Journal of Roentgenology*, 197(5):1122-1129
- Tanino, M., Takizawa, H., Yamamoto, S., Matsumoto, T., Tateno, Y. & Iinuma, T. (2003) A detection method of ground glass opacities in chest x-ray ct images using automatic clustering techniques. *Proceedings SPIE Medical Imaging 2003: Image Processing*, 5032:1728-1737
- Tang, J., and Guo, S. (2011) Segmentation of skin cancer using external force filtering snake based on wavelet diffusion. *Multi Modality Sate-of-the-Art Medical Image Segmentation and Registration Methodologies*. Springer. Berlin, 129-142.
- Tjondronegoro, D, Chen, YPP (2010) Knowledge-discounted Event Detection in Sports Video, *IEEE Transactions on Systems Managment & Cybernetics, Part A*. 40(5):1009-1024, 2010.
- Tourassi, GD., Delong, DM., and Floyd CE., (2006) A study on the computerized fractal analysis of architectural distortion in screening mammograms, *Physics in Medicine and Biology*, 51(5): 1299-1312
- Tourassi, GD., Harrawood, B., Singh, S., Lo, JY., and Floyd CE., (2007) Evaluation of information-theoretic similarity measures for content-based retrieval and detection of masses in mammograms, *Medical Physics*, 34(1): 140-150.
- Umbaugh, SE., Moss, R. H. & Stoecker, W. V. (1989) Automatic color segmentation of images with application to detection of variegated coloring in skin tumors. *IEEE Engineering in Medicine and Biology*

Magazine, 8(4): 43-50.

- Van Asselen, B., Schwarz, M., Van Vliet-Vroegindeweij, C., Lebesque, J. V., Mijnheer, B. J. & Damen, E. M. F. (2006) Intensity-modulated radiotherapy of breast cancer using direct aperture optimization. *Radiotherapy and Oncology*, 79(2): 162-169.
- Vecchiola, C., Pandey, S. and Buyya, R. (2009) High-Performance Cloud Computing: A View of Scientific Applications, *Proceedings 10th International Symposium on Pervasive Systems, Algorithms, and Networks*, 4-16
- Veldkamp, W. J. H. & Karssemeijer, N. (2000) Normalization of local contrast in mammograms. *IEEE Transactions on Medical Imaging*, 19(7):731-738.
- Verma, B. & Zakos, J. (2001) A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques., *IEEE Transactions on Information Technology in Biomedicine*, 5(1): 46-54.
- Vos, PC., Hambrock ,T., Hulsbergen-van de Kaa, CA., Futterer, JJ., Barentsz, JO., and Huisman, HJ., (2008) Computerized analysis of prostate lesions in the peripheral zone using dynamic contrast enhanced MRI. *Medical Physics* 35:888-899
- Wu, D., Erwin, D., and Rosner, GL. (2011) Sojourn time and lead time projection in cancer screening, *Lung Cancer*, 72(3):322-326.
- Ye, X., Lin, X., and Dehmeshki, J., (2009) Shape-based computer-aided detection of lung nodules in thoracic CT images, *IEEE Transaction on Biomedical Engineering*, 56:1810-1820.
- Yin Y., Fotin, SV., Periaswamy, S., Kunz, J., Haldankar, H., Muradyan, N., Turkbey, B., and Choyke, P., (2012) Fully automated 3D prostate central gland segmentation in MR images: a LOGISMOS based approach, *Proceedings SPIE Medical Imaging: Image Processing*, 8314:83143B
- Yuksel, M.E., and Borlu, M., (2009) Accurate Segmentation of Dermoscopic Images by Image Thresholding Based on Type-2 Fuzzy Logic, *IEEE Transactions on Fuzzy Systems*, 17(4): 976-982.
- Zhang, X. P. & Desai, M. D. (2001) Segmentation of bright targets using wavelets and adaptive thresholding. *IEEE Transactions on Image Processing*, 10(7):1020-1030.
- Zhao, B., Ginsberg, M. S., Lefkowitz, R. A., Jiang, L., Cooper, C. & Schwartz, L. H. (2004) Application of the ldm algorithm to identify small lung nodules on low-dose msct scans. *Proceeding SPIE Medical Imaging 2004: Image Processing*, 5370:818-823.
- Zheng, B., Lu, A., Hardesty, LA., Sumkin, JH., Hakim, CM., Ganott, MA., and Gur, D., (2006) A method to improve visual similarity of breast masses for an interactive computer aided diagnosis environment, *Medical Physics*, 33(1): 111-117.
- Zheng, Y. (2010), Breast cancer detection with Gabor features from digital mammograms, *algorithms*, 3(1): 44-62.

Highlights:

1. Studies and compare the recent works in different types of cancer detection
2. Low level features comparison for detecting different cancer types
3. Compare image modalities and associated segmentation algorithms
4. Research extension discussion in intermediate feature analysis and cloud structure for cancer detection



MOL2NET'22, Conference on Molecular, Biomedical & Computational Sciences and Engineering, 8th ed.

FROM MOLECULES TO NETWORKS



Computer Vision Based Skin Cancer Classification by Using Texture Features

Samreen Naeem ^{*,a}, Aqib Ali ^a and Sania Anam ^b

^a College of Automation, Southeast University, Nanjing, China.

^b Department of Computer Sci, Govt Associate College for Women Ahmadpur East, Bahawalpur, Pakistan.

* Corresponding author: samreencsit@gmail.com

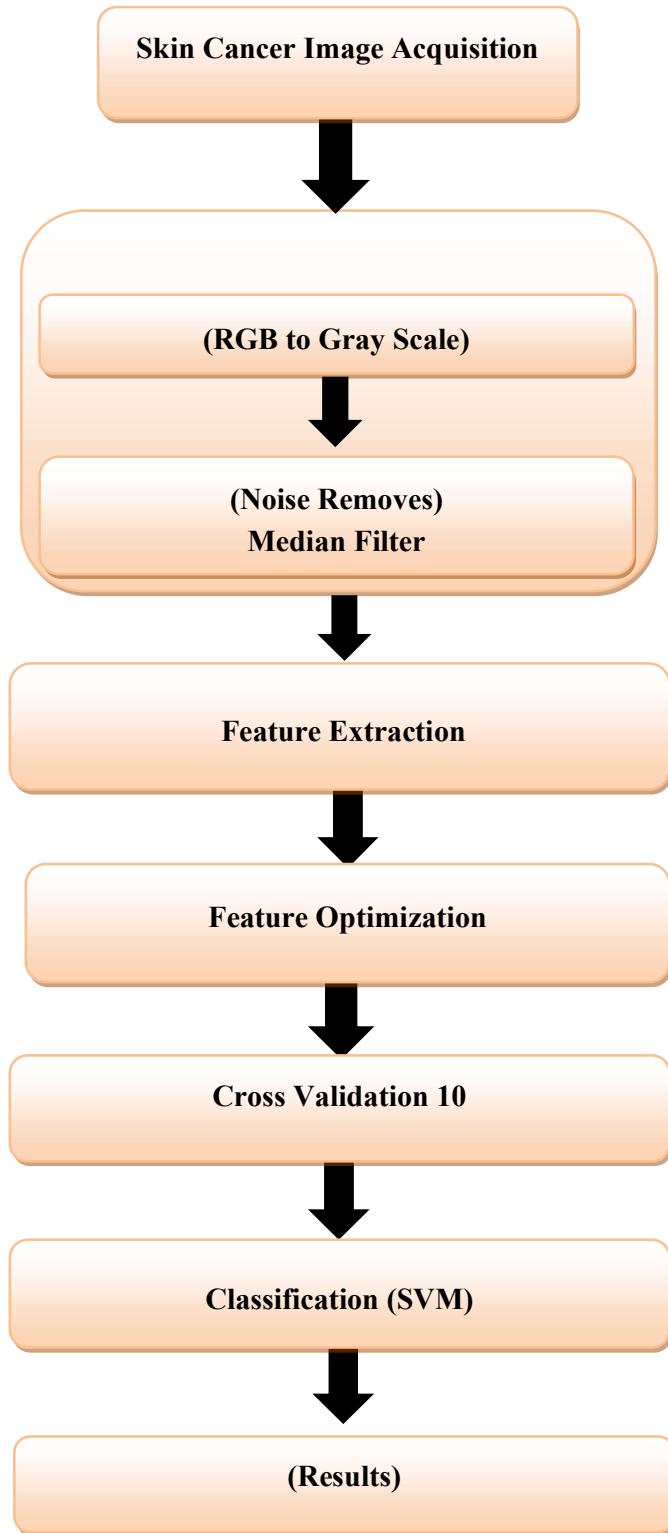
Abstract.

Cancer of the skin is now one of the most prevalent forms of the disease among people. As a result, accurate diagnosis of malignant lesions is of utmost significance in treating skin cancer. Dermatoscopic images can be used with computer-aided diagnostic tools, which may include machine learning models, to assist medical professionals in diagnosing skin cancer. In this particular research project, skin lesions were classified using image processing and machine learning strategies. Several distinct mathematical techniques have been implemented in the field of image processing in order to improve image quality. Image segmentation utilizing the watershed approach was conducted after an image preparation step, which included filtering the undesired pixels in the pictures. Following that, the lesioned regions were separated, and texture feature extraction was carried out. In the end, the classification was completed using the SVM algorithm, which stands for support vector machines. When the results acquired from the classifiers were compared, it was seen

that the SVM classifier had an accuracy of 94.33%.

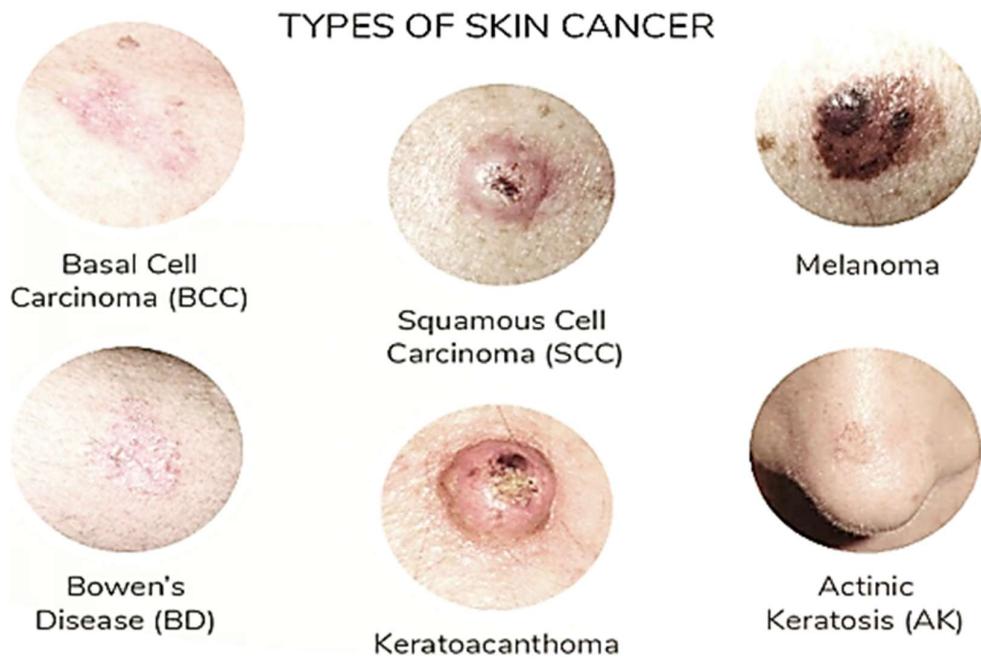
Keyword: Dermatoscopic Images, Skin Cancer, SVM, Machine Learning.

Graphical Abstract



Introduction

The growth and division of healthy cells result in the formation of new cells. These new cells will eventually take the place of injured or elderly cells. *Cancer* is the name of the disease that occurs when the body's normal cells multiply uncontrollably. Cancer cells, rather than killing off, continue to proliferate and give rise to more aberrant cells. A tumor can develop when there is aberrant cell proliferation in the epidermal layer, which may destroy other healthy tissues. This process, known as metastasis, describes how cancer cells move from one location in the body to another through the circulatory and lymphatic systems. Cancer comes in a wide variety of forms [1].



The most frequent types of cancer include thyroid cancer, lung cancer, breast cancer, bladder cancer, kidney cancer, colon and rectum cancer, leukemia, pancreatic cancer, prostate cancer, and skin cancer. Other less common types of cancer include leukemia and pancreatic cancer. The surface of our skin is the most significant organ in our bodies [2]. This organ's primary responsibilities include acting as a protective barrier against potentially hazardous substances that may enter the body from the environment, regulating body temperature via the hair, sweat glands, and adipose tissue that it contains, ensuring proper fluid and electrolyte balance, and contributing to the overall health of our bodies. The layers of our skin may be broken down into three broad categories. The epidermis, the topmost layer of our skin, is the first layer we will discuss [3].

In most cases, this layer stops the loss of fluid and serves as protection for the tissues that lie underneath it. The dermis is the second layer of the skin. In addition to hair follicles and sebaceous glands, the dermis is home to blood vessels and nerve fibers. It is in charge of regulating the

temperature and maintaining the fluid-electrolyte balance [4]. The hypodermis is the third and last layer of our skin. This layer is composed of adipose tissue. It not only guards against the damaging effects of impact on structures like bone and muscle but also helps maintain a steady body temperature. These three layers of our skin are fertile ground for the growth of tumors. The specific sort of cancer that an individual has depends on which tissues and layers of their skin the altered cells originate in. As a result, the procedure of treating the kind of cancer that was detected in the patient may be challenging or straightforward [5].

Skin cancer is the most prevalent form of the disease in every region of the globe surveyed by the Skin Cancer Foundation (SCF). According to the statistics from the Skin Cancer Foundation (SCF), more than two persons lose their lives to skin cancer every hour in the United States [6].

Materials and Methods

The ISIC 2020 training set, HAM10000, identified seven skin disorders in this investigation [7]. The dataset contains 300 150x2 RGB skin lesion photos. The significant number of lesions per category separates this dataset from others. These photos were classed as abnormal and usual. The dataset contains 300 malignant and 300 benign skin lesions. Preprocessing Ham10000 photos improved picture quality. Images received contrast-limited adaptive histogram equalization, morphological occlusion, and median filter. Image segmentation follows preprocessing. Image segmentation enhances meaning and analysis. Segmentation typically determines analysis success. Medical picture segmentation requires ROI extraction. Area-based segmentation uses the watershed transform to define the region of interest (ROI) and choose the ROI closest to the skin lesion.

Feature extraction is commonly used to classify segmented pictures in the training and test sets. This research extracted features using texture analysis [8]. Asymmetry is a crucial indicator of skin lesion malignancy. It measures how comparable the lesion form is along the primary axis. Vertical and horizontal asymmetry are computed independently. The compact index (CI) measures border irregularity as the ratio of the lesion circumference square to its area. Color is crucial to skin disease diagnosis. The normalized standard deviation of lesion red, green, and blue components measures color variance. The texture rule defines *diameter* as the most significant distance between any two sites of the lesion boundary. The lesion diameter is the diameter of the circle with the lesion area. Image texture determines pattern and color constancy. Haralick texture characterizes texture-based images by computing the gray-level co-occurrence matrix (GLCM). GLCM is used to extract texture features since it can calculate many characteristics and is easy to use. The co-occurrence matrix extracts 14 characteristic texture properties from the probability matrix. This research selected

Haralick's contrast, correlation, energy, and homogeneity [9]. GLCM is based on picture-pixel neighborhoods. It records the complete picture by searching for adjacent pairs of pixel values.

Correlation measures the combined probability of pixel pairings in each row and column. Energy, the square root of the total of square pixels, is the recurrence of pairs of pixels in the picture. Contrast distinguishes items by hue or color. Homogeneity; measures how near the GLCM component distribution is to its diagonal [10].

Results and Discussion

- Support vector machine (SVM)
- Time taken to build the model: 0.38 seconds
- Test mode: 10 fold

Table 1: SVM Classifier Summary

Total Number of Instances	300	
Correctly Classified Instances	283	94.3333 %
Incorrectly Classified Instances	17	5.6667 %
Kappa statistic	0.8867	
Mean absolute error	0.0777	
Root mean squared error	0.2294	
Relative absolute error	15.5438 %	
Root relative squared error	45.8745 %	

Table 2: SVM Classifier Detailed Accuracy

TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	Class
0.913	0.027	0.972	0.913	0.942	0.888	0.953	Normal
0.973	0.087	0.918	0.973	0.945	0.888	0.953	Abnormal
0.943	0.057	0.945	0.943	0.943	0.888	0.953	Weighted Avg.

Table 3: Confusion Matrix result using SVM Classifier

Classified as	A	B
A = Normal	137	13
B = Abnormal	4	146

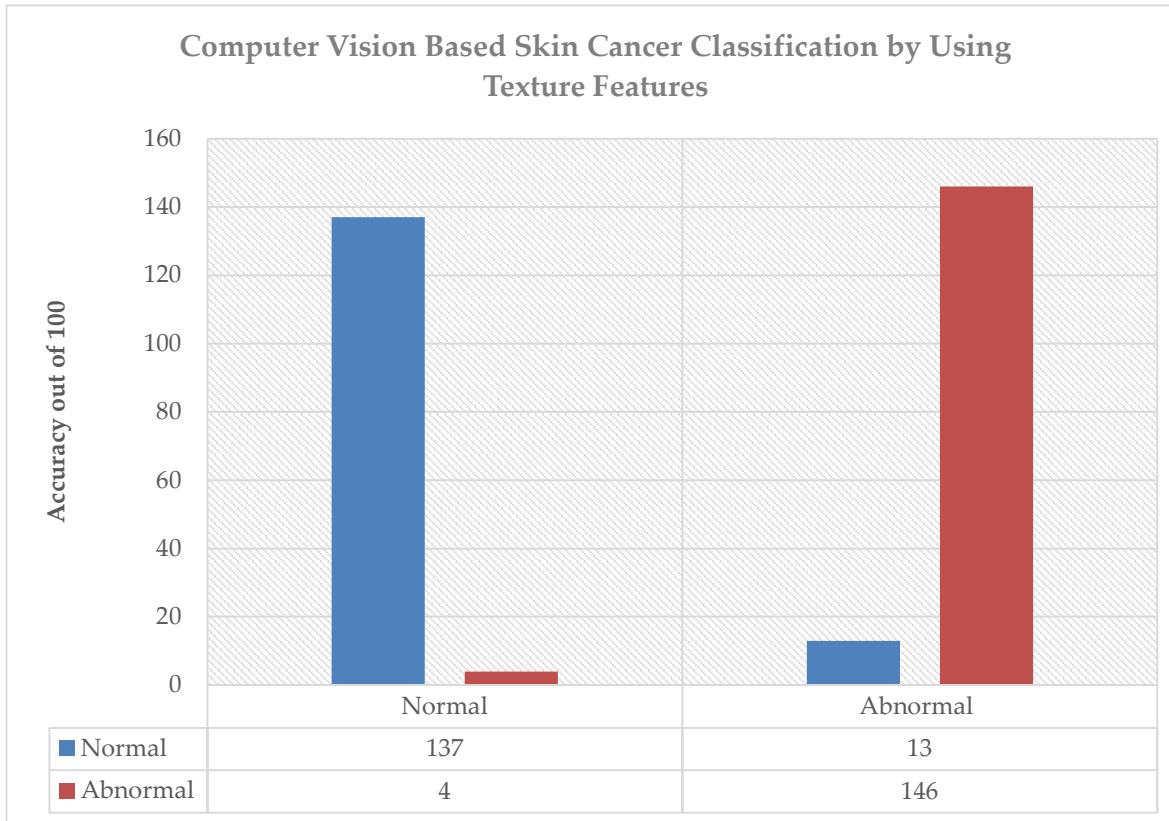


Figure 1: Accuracy of Dataset using SVM Classifier

Conclusions

This work uses machine learning to classify skin lesions in HAM10000 dermatoscopic pictures. SVM were used to classify the dataset's pictures after texture analysis segmented and extracted lesion characteristics. The research found that the DVM approach had greater accuracy, precision, and f1-score. Combining an optimized classifier, a current segmentation technique, and a well-tested feature extraction strategy shows SVM classifier performance. This study's findings may be improved by increasing training samples and balancing labeled lesion groups. Using sophisticated classification models or deep learning techniques and training the models on the complete data set may also improve outcomes.

References

- [1]. Leiter, U., Keim, U., & Garbe, C. (2020). Epidemiology of skin cancer: update 2019. Sunlight, Vitamin D and Skin Cancer, 123-139.
- [2]. Neagu, M., Constantin, C., Cretoiu, S. M., & Zurac, S. (2020). miRNAs in the Diagnosis and Prognosis of Skin Cancer. Frontiers in cell and developmental biology, 8, 71.
- [3]. Cullen, J. K., Simmons, J. L., Parsons, P. G., & Boyle, G. M. (2020). Topical treatments for skin cancer. Advanced drug delivery reviews, 153, 54-64.
- [4]. Tschandl, P., Rinner, C., Apalla, Z., Argenziano, G., Codella, N., Halpern, A., ... & Kittler, H. (2020). Human–computer collaboration for skin cancer recognition. Nature Medicine, 26(8), 1229-1234.
- [5]. Zhang, N., Cai, Y. X., Wang, Y. Y., Tian, Y. T., Wang, X. L., & Badami, B. (2020). Skin cancer diagnosis based on optimized convolutional neural network. Artificial intelligence in medicine, 102, 101756.
- [6]. Ilkhomovna, K. D. (2021). Modern Look of Facial Skin Cancer. BARQARORLIK VA YETAKCHI TADQIQOTLAR ONLAYN ILMIY JURNALI, 1(1), 85-89.
- [7]. Razmjooy, N., Ashourian, M., Karimifard, M., Estrela, V. V., Loschi, H. J., Do Nascimento, D., ... & Vishnevski, M. (2020). Computer-aided diagnosis of skin cancer: a review. Current Medical Imaging, 16(7), 781-793.
- [8]. Ali, A., Mashwani, W. K., Naeem, S., Uddin, M. I., Kumam, W., Kumam, P., ... & Chesneau, C. (2021). COVID-19 infected lung computed tomography segmentation and supervised classification approach.
- [9]. Naeem, S., Ali, A., Qadri, S., Khan Mashwani, W., Tairan, N., Shah, H., ... & Anam, S. (2020). Machine-learning based hybrid-feature analysis for liver cancer classification using fused (MR and CT) images. Applied Sciences, 10(9), 3134.
- [10]. Ali, A., Qadri, S., Khan Mashwani, W., Kumam, W., Kumam, P., Naeem, S., ... & Sulaiman, M. (2020). Machine learning based automated segmentation and hybrid feature analysis for diabetic retinopathy classification using fundus image. Entropy, 22(5), 567.