Utilising Artificial Intelligence to Augment Physiotherapy Education: A Video-Based Analysis and Grading System for Evaluating Student Therapists' Performance

Victor Umesiobi
Teesside University,
Middlesbrough, United Kingdom.
q2094871@tees.ac.uk.

Hiba Alsmadi
Teesside University,
Middlesbrough, United Kingdom.
h.alsmadi@tees.ac.uk.

Victor Hammond
Teesside University,
Middlesbrough, United Kingdom.
c2092063@tees.ac.uk.

Gok Kandasamy
Teesside University,
Middlesbrough, United Kingdom.
g.kandasamy@tees.ac.uk.

Ala S. Al Kafri
Teesside University,
Middlesbrough, United Kingdom.
a.alkafri@tees.ac.uk.

2024 17th International Conference on Development in eSystem Engineering (DeSE) | 979-8-3503-6869-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/DeSE63988.2024.10911886

***Abstract*— The rapid evolution of artificial intelligence (AI) has ushered in new opportunities for enhancing physiotherapy education. This study introduces an innovative AI-based tool that integrates computer vision and speech recognition to provide a comprehensive and objective assessment of physiotherapy students' practical skills. Designed to surpass the limitations of traditional, subjective methods, this system analyses both physical movements and verbal interactions during exams, categorizing student performance into 'Good', 'Average', or 'Brief' to offer a multifaceted evaluation of their clinical abilities. The research aimed to develop and validate a system that accurately captures and analyses nuanced physical and verbal behaviours in physiotherapy practical exams. The system was trained and tested on a curated dataset of recorded sessions, where it demonstrated high precision, achieving an overall accuracy of 89% in identifying specific actions performed by students. These results underscore the system's capability to enhance the objectivity and depth of practical assessments significantly. This study marks a significant contribution to healthcare education by showing the effectiveness of AI technologies in assessing practical and communication skills. The AI-assisted tool not only improves the educational process but also better prepares students for clinical practice. The findings suggest a model for future AI integration in professional healthcare training and assessment, highlighting the transformative potential of AI in advancing physiotherapy education.**

Keywords—Physiotherapy Education, AI-based Assessment Tools, Physiotherapy Practical Skills Assessment, Objective Assessment Methodologies, Technology in Education

I. Introduction

The integration of artificial intelligence (AI) and computer vision technologies in physiotherapy education marks a significant shift from traditional teaching methods to a more advanced, technology-focused approach. This evolution emphasizes digital literacy and technological proficiency to improve the learning experience and prepare students for contemporary clinical practice [1][2].

Historically, physiotherapy education has prioritized practical skills and direct patient interaction, focusing on manual therapy and tailored treatment strategies. However, traditional educational and evaluation methods have shown limitations in scalability, objectivity, and comprehensiveness, often leading to inconsistencies in student outcomes and raising concerns about the reliability and fairness of these approaches [3].

The adoption of AI in physiotherapy education has introduced tools that can objectively assess student performance, provide immediate feedback, and enhance understanding of complex clinical situations through simulations and virtual reality [4]. AI technologies, including machine learning and computer vision, enable precise assessment of student techniques in practical exams, surpassing the capabilities of human evaluators.

AI also enriches the educational journey beyond assessment, offering customized learning paths, interactive simulations, and predictive analytics to gauge student performance and engagement. This utilization of AI supports personalized, flexible, and engaging educational experiences, aligning with modern educational theories that focus on learner-centred approaches, feedback, and reflective practice [5].

While AI holds promise for enhancing physiotherapy education, ethical issues, data protection, and the necessity for robust validation procedures must be addressed to ensure the tools' effectiveness and accuracy [6]. Integrating such technologies requires careful consideration of innovation and ethical responsibility, with collaboration among educators, developers, and policymakers to establish protective standards and guidelines.

Amid challenges, the future of physiotherapy education leans towards digital and AI-enhanced methods, accelerated by the COVID-19 pandemic's push for remote learning technologies and the need for digital tools to sustain educational quality under constraints [7]. The expanding use of AI in curricula addresses educational challenges and reshapes healthcare education in the digital age.

II. Problem Statement

The assessment of practical skills in physiotherapy education is challenged by traditional, subjective methods that often lead to inconsistencies and biases. These techniques, reliant on educator judgment, fail to fully assess students' capabilities, particularly in technique execution and patient communication. The lack of objective metrics and advanced tools for evaluating verbal interactions during exams limits precise feedback and hinders student development. This project proposes an innovative system that integrates computer vision and speech recognition technologies to enhance the assessment of physiotherapy students' practical skills and communication abilities.

III. Research Objectives

This research project explores integrating video analysis and voice recognition technologies to improve assessment methods for physiotherapy students during practical exams. It aims for a detailed and unbiased evaluation of students' practical skills and verbal interactions. The project will employ video analysis to detect and track students' actions in simulated exams, enhancing the accuracy of performance assessments. Voice recognition will evaluate students' communication with patients, focusing on terminology use and clarity. The integration of these technologies into a single assessment tool marks a significant advancement in physiotherapy education, aiming for fair and thorough evaluations based on actual student performance.

The research seeks to develop a framework that encapsulates both the physical and verbal dimensions of physiotherapy, providing a more precise and objective assessment. Focus areas include:

- Conceptualizing a computer vision system for analyzing physiotherapy movements during exams.
- Utilizing motion detection algorithms to accurately track and evaluate therapists' techniques.
- Improving educators' capabilities in assessing students' practical and communicative skills comprehensively.
- Integrating speech recognition technology to evaluate verbal interactions with patients.

This method aims to ensure fair and precise evaluations, promoting better education in physiotherapy through technology integration.

IV. Systematic Review of Related Work

The integration of computer vision and speech recognition is transforming physiotherapy education, moving assessments towards objective and comprehensive evaluations. These technologies improve exam accuracy and student learning, preparing them for healthcare roles. The review addresses three primary areas: the application of computer vision in practical exams, speech recognition for evaluating communication skills, and overcoming traditional method limitations.

This synthesis highlights significant research, outlining the potential and challenges of these technologies in enhancing physiotherapy education. It underscores the need for innovative assessment approaches that align with the complexities of modern physiotherapy practice.

A. The Role of Computer Vision in Enhancing Practical Examinations

Computer vision, part of artificial intelligence, helps computers interpret visual information like humans, enhancing physiotherapy education by providing precise analyses of students' clinical skills [8]. Zhang and Hossain (2019) demonstrated its potential to improve skill assessment accuracy by identifying procedural executions and offering detailed feedback beyond traditional methods [8]. Esteva et al. (2017) further validated its reliability by using deep learning to classify skin cancer with high accuracy [9]. Ker et al. (2018) noted its effectiveness in assessing physical movements and postures in physiotherapy [10]. Challenges include the need for large datasets for effective training, yet the potential for improved assessment accuracy remains significant, argued by Tajbakhsh et al. (2016) [12]. Computer vision also allows for real-time feedback, significantly enhancing the learning process by enabling immediate improvements [4].

B. Speech Recognition: Assessing Communication Skills in Physiotherapy Education

Patel & Hossain (2018) highlighted speech recognition's role in enhancing clinical communication by objectively analyzing medical terminology and interaction clarity in physiotherapy assessments [13]. McGinnis et al. (2016) emphasized its capacity to measure crucial soft skills like effective communication [14]. Despite the potential benefits, challenges like user resistance and the need for customization to specific clinical settings persist, as noted by Blackley et al. (2019) [15]. Korstjens & Moser (2018) and Oviatt (2012) discussed the technology's credibility and the importance of user-friendly interfaces for effective educational application [16][17]. Speech recognition could revolutionize communication skill assessments in physiotherapy education if these hurdles are overcome.

C. Addressing Traditional Assessment Limitations

Traditional physiotherapy assessments often rely on subjective evaluations by educators, leading to variability in outcomes, as noted by Turner and Foster (2019) and Dalton et al. (2011), who pointed out that these methods might not fully capture the complexity of required clinical skills [18][19]. Harris and Smith (2021) suggested using computer vision to objectively analyze movements during exams, providing a quantifiable performance measure [4]. Similarly, speech recognition, as discussed by Patel and Hossain (2018), could offer an objective way to assess communication skills [13]. Despite the potential of these technologies to enhance fairness and objectivity in assessments, challenges like ethical concerns and the need for validation remain significant, as Aebersold (2018) highlighted [20].

V. Methodology

The methodology section outlines the development of an AI-based tool to evaluate healthcare students, specifically in physiotherapy. It details strategies and technologies to assess students' practical skills and communication during assessments, using computer vision and voice recognition to shift from subjective to objective evaluations. The section systematically presents the experimental design, data collection methods, and analysis techniques, structured into key areas like research design, technology exploration, and ethical considerations. This approach aims to enhance the objectivity and precision of student evaluations in physiotherapy education.

A. Research Design

This research designs a theoretical framework for an AI system to better evaluate physiotherapy students' skills and communication. Initial steps involved a comprehensive literature review across platforms like PubMed and IEEE Xplore, using relevant keywords, which informed the development of the theoretical framework. Despite reviewing over 115 publications, the integration of computer vision and voice recognition in physiotherapy assessments remained unaddressed, leading to the selection of 25 pertinent articles. The project progresses from model building to testing with simulated data, ensuring ethical data use, and exploring potential scalability and data representativeness issues.

B. Technology Exploration

This section focuses on employing computer vision and voice recognition to revolutionize physiotherapy student evaluations. Computer vision processes visual data to provide consistent, unbiased assessments of physical techniques, while voice recognition technology assesses communication skills by analyzing audio recordings of student-patient interactions. The integration of these technologies is vital for a holistic assessment system that effectively combines physical and communicative skills.

C. Data Collection and Preprocessing

1) Data sources and acquisition

The initial stage of data collecting and preprocessing is the careful selection and gathering of video data from YouTube. The platform offers a comprehensive dataset, which contains videos of a wide range of physiotherapy practical assessments. These videos are carefully selected under the supervision of qualified physiotherapists and physiotherapy education instructors to appropriately reflect scenarios that students may encounter during evaluations. The recordings incorporate both visible motions and vocal conversations, providing a complete picture of a student's capabilities and communication abilities, which are essential components for a thorough assessment in physiotherapy training.

2) Video data annotation

Following the acquisition of video data, the project uses Supervisely [21], a web-based tool with comprehensive video annotation capabilities. This platform allows frame-by-frame analysis without converting videos into still images, crucial for maintaining the continuity of physiotherapy movements. The tool handles large video files and features a timeline panel for an efficient overview of the video's annotation status, facilitating segment identification and tagging adjustments.

The dataset includes 44,361 video frames in high definition (1280x720 pixels), annotated to capture detailed physiotherapy assessments. Annotations identify different postural assessments and therapist-patient interactions, using 34 action classes to categorize physiotherapy examination actions, ensuring consistency and accuracy across the dataset.

Key annotation classes include:

- Postural Landmarks and Examinations: Identifying skeletal structure alignment.
- Dynamic Assessments: Assessing body part functions like lumbar movements and thoracic rotation.
- Patient Interaction: Evaluating communication aspects during therapy.
- General Observations: Capturing professionalism and setting context.

Each frame is meticulously labelled with bounding boxes that define spatial locations for actions and interactions, such as a knee examination. As seen in Figure 1, this approach, focusing on broader areas, effectively assesses postures and alignments. Supervisely enhances this process with AI technologies like MixFormer for bounding box detection and other tools for mask and polygon segmentation, improving the labelling speed and accuracy.

Fig 1. A sequence of annotated video frames displaying a physiotherapy session, with red bounding boxes indicating key postural assessment for motion tracking and analysis using the Slow Fast network model.

These annotations undergo a rigorous review by physiotherapy experts to meet the exact standards necessary for AI model training, ensuring precise and consistent dataset annotations.

3) Converting to Structured Data Formats

Once annotated, the data is compiled into JSON format, which Supervisely supports natively. This format offers precise descriptions of video frames, annotations, and essential information. It also contains detailed descriptions of video frames, annotations, and other metadata necessary for AI processing. These JSON files are then transformed into CSV format, which organizes the data into a tabular structure to match the format of the AVA dataset [22], a large-scale benchmark for atomic visual actions, suited for machine learning algorithms. The CSV files include extensive annotation information, such as the video frames, file paths, bounding box coordinates, class names, and other relevant features required for AI model training.

4) Contribution to AI model development

The organized dataset, now in CSV format, serves as the foundation for training computer vision and speech recognition algorithms. These models are designed to objectively evaluate physiotherapy students in simulated exam settings, identifying and assessing the technical precision of physical treatments and the quality of student-patient communication. This comprehensive approach to data collecting and preprocessing assures that the AI model generated is reliable and accurate, paving the way for the transformation of physiotherapy education. This phase improves learning outcomes and prepares future practitioners by offering tools for objective, thorough, and scalable assessments of student performance.

VI. Model Development

In physiotherapy education, accurately evaluating students' practical exams is essential. This project leverages the SlowFast network and OpenAI's Whisper model to analyze physiotherapy sessions in detail. Utilizing the PyTorch-based SlowFast framework, the model detects physiotherapy actions and excels in motion recognition [23][24]. The Whisper model enhances this by providing detailed speech and audio analysis, crucial for understanding spoken interactions [25]. This integration aims to improve assessment precision and educational quality by delivering thorough feedback on technique performance, objectively assessing student skills, and identifying areas for improvement. The SlowFast action detection model is meticulously engineered to analyze physical motions in video data with high efficiency on NVIDIA RTX-A6000 GPUs. It utilizes the SlowFast networks from Facebook AI Research's PyTorch video model repository, optimally leveraging its deep learning capabilities over 100 training epochs for physiotherapy motion detection [23]. The model processes CSV-formatted physiotherapy data featuring frame-wise annotations and metadata. To enhance resilience against visual variations, video frames undergo normalization, resizing to 250 pixels, RGB colour conversion, and various augmentation techniques, including rotation and flipping. The network architecture features a Slow route with deeper layers for detailed spatial information, and a Fast route for capturing rapid movements, both critical in physiotherapy analysis.

Using advanced optimizers like Adam with momentum, alongside a loss function mix of cross-entropy and smooth L1, the model effectively identifies and localizes physiotherapy actions, ensuring robust action detection optimized for short training durations and high throughput as shown in Figure 2.

1) Performance Categorization and Grading

The evaluation of physiotherapy student performance began by establishing a comprehensive standard, derived from historically assessed video submissions by physiotherapy educators, categorized into three performance levels: 'Good', 'Average', and 'Brief'. This formed the basis for a trained action detection model that recognized and categorized actions in new video submissions, organized into corresponding folders to build a pre-processed dataset, which included action sequences and grading labels for accurate subsequent assessments.

Temporal accuracy was achieved by marking the start and end times of each action and discussion, producing a detailed, synchronized report of performance. The DeLLMa (Decision-making Large Language Model Assistant) framework, as proposed by Liu et al. (2023) [26], utilized decision and utility theories to refine grading predictions. It featured a custom grading agent that assessed various performance dimensions such as 'action completeness', 'discussion quality', and 'actions-discussion alignment'. This approach allowed for a multi-dimensional evaluation of student competencies, enhancing grading precision through a multi-step scaffolding method aimed at maximizing accuracy by accounting for uncertainties in student performances, thus ensuring a robust assessment compatible with the complexities of physiotherapy education.

2) Model Architecture

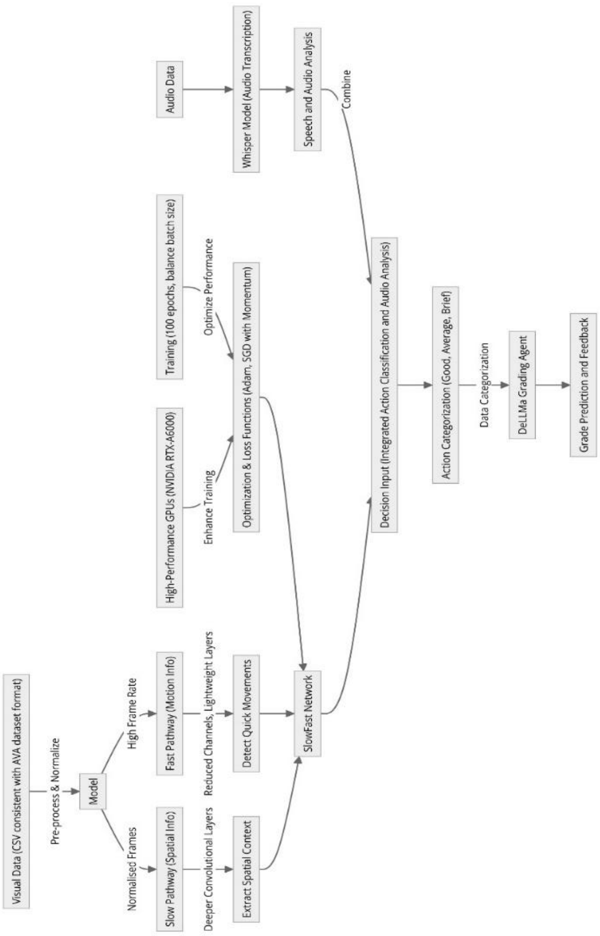

Fig2. Schematic representations of the integrated model architecture combining SlowFast Network and Whisper Model for physiotherapy education assessment.

VII. EXPERIMENTATION AND TESTING

Following the formulation of performance criteria and data categorization, the next part of the study technique included rigorous experimentation and testing to evaluate the integrated model's efficacy. This experiment was created to test the hypothesis that the model could accurately mimic the grading supplied by experienced educators based on an analysis of the student's actions and conversations in the video submissions.

The SlowFast networks examined the students' motions, capturing the nuances of their physical evaluations, while OpenAI's Whisper carefully evaluated the aural components of the student's interactions with patients. The DeLLMa framework, embedded within this system, then synthesized these observations to predict grading outcomes. The experiment required uploading each video to the API and letting the model complete its analysis. Actions and discussions were extracted and categorized as 'Good', 'Average', and 'Brief'. The model rated each student's performance using the subtle probabilistic evaluations enabled by the DeLLMa architecture. To assess the model's accuracy, the grading predictions were compared to historical grades provided by educators an example is shown in Figure 3.

The purpose of these experiments was to illustrate the potential for AI-assisted grading systems to supplement and, in some circumstances, improve existing ways of evaluating student performance in clinical education settings. The integration of AI technologies, specifically computer vision and speech recognition, in the assessment of physiotherapy students, has proven highly effective, significantly enhancing the objectivity and accuracy of evaluations. This study's AI-enhanced model, incorporating the DeLLMa framework, action detection, and audio analysis, has demonstrated a notable advancement in the method and quality of student assessments. The multimodal approach not only expedited the grading process but also ensured impartiality and consistency, attributes that are often challenging to maintain with manual assessments.

VIII. Results and Discussion

The experimental phase yielded impressive results: with an accuracy rate of 89%, the model aligned remarkably well with the grading standards of experienced physiotherapy educators, indicating its sophisticated capacity to discern and evaluate the intricate aspects of clinical performance integral to physiotherapy education. This high degree of accuracy reflects the model's effectiveness in managing the complex variations found in practical physiotherapy examinations. Moreover, the recorded test loss of 0.6259 highlights the model's precision in quantifying student performance, further underlining its capability to differentiate between the various levels of student demonstrations. Overall, the implementation of AI technologies in physiotherapy programs not only enhances the evaluation process but also impacts teaching and learning significantly. With accurate data on student performance, educators can tailor educational strategies to meet individual and collective needs, providing immediate, detailed feedback that accelerates learning curves and improves skill acquisition. This fosters a more engaged and responsive educational environment, marking a significant stride forward in the domain of skill-based education and setting a new standard for the integration of AI in healthcare professional training and assessment.

IX. Future Research Development and Limitations

Transitioning from open-source datasets to real student exam data will improve the accuracy of AI applications, allowing them to better address specific educational needs. However, several limitations must be considered: Evaluation Complexity: Evaluating student performance can be challenging, as it requires assessing actions separately from the broader decision-making process. This complexity can make it difficult to determine the true effectiveness of AI interventions. Processing Speed: The model's performance may be hindered by its need to process individual actions before combining them into a cohesive context.

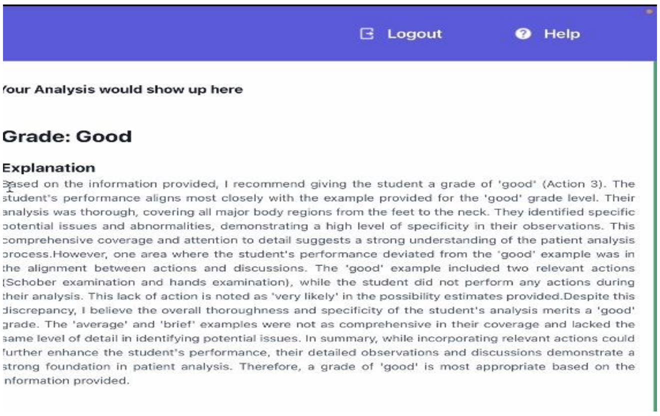

Fig 3. Screenshot of a classification test interface for evaluating physiotherapy student performances, displaying a successful request with detailed grading decision in the response. The UI confirms the 'Good' grade based on the video analysis and demonstrates the model's assessment capabilities.

This step-by-step approach can slow down real-time analysis, impacting the responsiveness of AI tools during assessments. Data Sufficiency: Given the intricate context in which large language models (LLMs) through DeLLMA operate, there may not be enough data to support optimal decision-making. This limitation could affect the model's ability to provide accurate feedback or guidance, especially in nuanced situations. Real-Time Data Analysis: Future research should focus on developing tools for real-time data analysis that consider factors such as stress levels and patient responses. However, implementing these tools poses challenges in ensuring their accuracy and reliability. Partnerships and Curriculum Integration: Collaborating with educational institutions is vital for effectively incorporating AI into curricula. Continuous feedback from users will be essential for refining AI applications, yet establishing these partnerships can be time-consuming and resource-intensive. Data Privacy and Ethical Standards: Ensuring data privacy in student assessments is paramount. Implementing strong ethical standards is necessary to protect sensitive information, but this can complicate data collection and usage. Longitudinal Studies: To assess the long-term impact of AI on student readiness and educational outcomes, longitudinal studies are required. These studies will help determine whether AI integration meets the evolving demands of healthcare education. By addressing these limitations and focusing on adaptive systems that utilize real-time data, the integration of AI in physiotherapy education can lead to more personalized learning experiences and improved skill development.

X. Conclusion

Research into AI technologies such as computer vision and speech recognition in physiotherapy education has displayed their potential to significantly enhance the evaluation of practical skills and communication abilities, thereby revolutionizing traditional assessment methodologies. These technologies can improve the fairness, accuracy, and comprehensiveness of assessments, which have traditionally relied on subjective methods prone to inconsistencies. The integration of AI allows for precise, real-time examination of technical skills and introduces a novel dimension to the analysis of verbal communications, which are crucial in-patient interactions. Addressing these challenges requires a collaborative effort among researchers, educators, technology developers, and policymakers, as well as continuous testing and adaptation of AI technologies. This collaboration is crucial to overcoming the barriers to AI integration and ensuring that these technologies meet the evolving demands of physiotherapy education.

As we continue to explore and expand the capabilities of AI in this field, this is just the beginning of a broader movement towards a more technologically advanced educational environment. Continued research, development, and ethical considerations are necessary to fully realize the potential of AI in enhancing educational outcomes and preparing students for the complexities of modern healthcare practice. This study sets the stage for future advancements, encouraging ongoing innovation and adherence to ethical standards to optimize the benefits of AI in physiotherapy education.

REFERENCES

[1] M.A. Jones and D.A. Rivett, "The Evolution of Physiotherapy: From Massage to Evidence-Based Practice," Physiother. Theory Pract., vol. 37, no. 1, pp. 12–22, 2021.

[2] J. Smith and S.E. Pivko, "Historical perspectives on physiotherapy education: From apprentice model to evidence-based practice," Int. J. Ther. Rehabil., vol. 27, no. 6, 2020.

[3] P. Turner and M. Foster, "Integrating evidence-based practice in physiotherapy education: Bridging the gap between academic training and real-world application," J. Physiother. Educ., vol. 33, no. 3, pp. 237–245, 2019.

[4] L. Harris and L. Smith, "Enhancing physiotherapy education with virtual reality and simulation-based learning: A review of recent developments," J. Physiother. Technol., vol. 5, no. 1, pp. 22–29, 2021.

[5] V. González-Calatayud, P. Prendes-Espinosa, and R. Roig-Vila, "Artificial Intelligence for Student Assessment: A Systematic Review," Appl. Sci., vol. 11, no. 12, art. no. 5467, 2021. DOI: 10.3390/app11125467.

[6] M. Ran, D. Banes, and M.J. Scherer, "Basic principles for the development of an AI-based tool for assistive technology decision making," Disabil. Rehabil. Assist. Technol., vol. 2020, DOI: 10.1080/17483107.2020.1817163, 2020.

[7] R. Martin and A. Bhatti, "Adapting physiotherapy education to remote learning in the wake of COVID-19," J. Online Learn. Health Sci., vol. 1, no. 2, pp. 34–42, 2020.

[8] Y. Zhang and M.A. Hossain, "Applications of Computer Vision in Medical Education and Diagnostic Radiology: A Systematic Review," J. Digit. Imaging, vol. 32, pp. 283–293, 2019.

[9] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, pp. 115–118, 2017.

[10] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," IEEE Access, vol. 6, pp. 9375–9389, 2018.

[11] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," Med. Image Anal., vol. 42, pp. 60–88, 2017.

[12] N. Tajbakhsh, J.Y. Shin, S.R. Gurudu, R.T. Hurst, C.B. Kendall, M.B. Gotway, and J. Liang, "Convolutional neural networks for medical image analysis: full training or fine-tuning?" IEEE Trans. Med. Imaging, vol. 35, no. 5, pp. 1299–1312, 2016.

[13] R. Patel and A. Hossain, "Improving clinical communication through structured conversations and speech recognition technology," J. Healthcare Eng., 2018, Article ID 4817965.

[14] P. Q. McGinnis, R. K. Guenther, and S. F. Wainwright, "Development and integration of professional core values among practising clinicians," Phys. Ther., vol. 96, no. 9, pp. 1417–1429, 2016.

[15] S. V. Blackley, D. T. Huynh, L. Wang, Z. T. Korach, and L. Zhou, "Speech recognition for clinical documentation from 1990 to 2018: A systematic review," J. Am. Med. Inform. Assoc., vol. 26, no. 4, pp. 324–338, 2019.

[16] I. Korstjens and A. Moser, "Series: Practical guidance for qualitative research. Part 4: Trustworthiness and Publishing," Eur. J. Gen. Pract., vol. 24, no. 1, pp. 120–124, 2018.

[17] S. Oviatt, "Human-centred design meets cognitive load theory: designing interfaces that help people think," in Proc. 14th ACM Int. Conf. Multimodal Interact., pp. 89–96, 2012.

[18] P. Turner and M. Foster, "Integrating evidence-based practice in physiotherapy education: Bridging the gap between academic training and real-world application," *J. Physiother. Educ.*, vol. 33, no. 3, pp. 237–245, 2019.

[19] M. Dalton, M. Davidson, and J. Keating, "The assessment of physiotherapy practice (APP) is a valid measure of the professional competence of physiotherapy students: A cross-sectional study with Rasch analysis," J. Physiother., vol. 57, no. 4, pp. 239–246, 2011.

[20] M. Aebersold, "Simulation-based learning is no longer a novelty in undergraduate education," Online J. Issues Nurs., vol. 23, no. 2, 2018.

[21] Supervisely, "Supervisely web-based annotation tool," 2023. [Online]. Available: https://app.supervisely.com/. [Accessed: 15 Mar. 2024].

[22] C. Gu, C. Sun, D. A. Ross, C. Vondrick, C. Pantofaru, Y. Li, et al., "Ava: A video dataset of spatio-temporally localized atomic visual actions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* (CVPR), pp. 6047-6056, 2018.R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[23] C. Feichtenhofer, H. Fan, J. Malik, and K. He, "SlowFast networks for video recognition," in Proc. IEEE/CVF Int. Conf. Comput. Vis., pp. 6202–6211, 2019.

[24] C. Feichtenhofer, "PyTorch video models," 2019. [Online]. Available: https://github.com/facebookresearch/SlowFast. [Accessed: 25 Feb. 2024].

[25] Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, et al., "Robust Speech Recognition via Large-Scale Weak Supervision," OpenAI Blog, 2022. [Online]. Available:

https://openai.com/blog/whisper/. [Accessed: 12 Feb. 2024].

[26] O. Liu, D. Fu, D. Yogatama, and W. Neiswanger, "DeLLMa: A Framework for Decision Making Under Uncertainty with Large Language Models," arXiv preprint arXiv:2305.11771, 2023.