Revisor 1:

**- Comentários e sugestões para autores**

O artigo sobre a avaliação da tomada de decisão assistida por IA por enfermeiros no tratamento de úlceras venosas sugere a necessidade de aumentar o tamanho da amostra de enfermeiros, além de informações adicionais sobre exames de imagem de úlceras venosas. Além disso, o diagnóstico e o tratamento de úlceras venosas exigem mais pesquisas sobre o papel dos médicos, além dos neurocirurgiões e enfermeiros.

**Linha 25. "Esta rede foi**

então utilizada por 25 enfermeiros através do aplicativo VenoTEC." . O tamanho da amostra é pequeno, com apenas 13 enfermeiros participantes. Provavelmente serão necessários enfermeiros adicionais.

**Linha 31.**

As legendas das imagens são obrigatórias, assim como imagens reais de pacientes com úlceras venosas.

**Linha 250.**

Embora a matriz de confusão seja representada graficamente, parece necessário criar uma tabela mostrando a especificidade, sensibilidade e recall de acordo com as estatísticas para ajudar na compreensão.

**Linha 277.**

Embora o diagnóstico e o tratamento de úlceras venosas dependam de redes neurais para apoiar as decisões dos enfermeiros, é necessário revisar mais a fundo o papel dos médicos no diagnóstico e tratamento da doença e comparar as correlações com o aplicativo VenoTec.

**Revisor 2**

Comentários e sugestões para autores

Caros autores,

Inicialmente, gostaria de expressar meu apreço pela ideia e esforço investidos na condução da pesquisa e na redação do manuscrito que me foi recomendado para revisão.

**Comentários gerais:**

* O trabalho utiliza um modelo RUP estruturado, que garante um fluxo de projeto lógico e consistente (fases de conceito, implementação e implantação). O uso de um ciclo de desenvolvimento bem desenvolvido (competitivo com os ensaios clínicos tradicionais) demonstra uma abordagem metodológica sólida e uma descrição clara das etapas subsequentes do trabalho.
* Os autores envolveram enfermeiros experientes no estudo, o que reforça a natureza prática do estudo. Esses especialistas participaram da definição das características clínicas das feridas e da sugestão de terapias, o que confere credibilidade à seleção dos dados de entrada para o modelo.
* O projeto de pesquisa foi aprovado pelo Comitê de Ética da Universidade do Estado do Rio Grande do Norte (CEUA-UFRN, nº 53258321.2.0000.5537). A presença dessa declaração (e do documento comprobatório pertinente) atesta a conformidade com os padrões de pesquisa médica e gera confiança na transparência dos procedimentos.

**Comentários específicos:**

O trabalho é apropriado para o periódico, mas há algumas observações que devem ser abordadas antes que o manuscrito possa ser aceito.

* O artigo não descreve a verificação independente da rede em um conjunto de dados separado ou em ambientes clínicos. Sem essa validação, é impossível avaliar o desempenho do modelo fora de um testador interno limitado. Testes em um grupo maior e diversificado de casos são necessários para confirmar sua utilidade prática.
* O resultado relatado de 100% de precisão nos dados de teste é suspeitosamente alto. Na prática de TI, alcançar uma classificação perfeita com um conjunto tão limitado significa que o modelo provavelmente será sobrecarregado. Os autores devem revisar o método de avaliação do modelo (por exemplo, introduzir validação cruzada ou um conjunto de testes separado) para confirmar as reais capacidades da rede.
* A descrição da rede sugere um modelo sequencial com uma camada de entrada contendo 5 características, duas camadas ocultas e uma camada de saída com 16 neurônios. No entanto, não há descrição das camadas convolucionais características das CNNs, o que levanta dúvidas sobre o uso real de imagens de feridas pelo modelo. Parece que a rede se baseia em características inseridas manualmente, em vez de análise automática de pixels. Uma representação tão imprecisa levanta dúvidas sobre sua compatibilidade com o nome "CNN".

O trabalho apresentado apresenta uma ideia interessante e inovadora de combinar inteligência artificial com a prática de enfermagem no tratamento de feridas venosas. Por um lado, aprecio a estrutura clara do projeto, sua confiança em conhecimento especializado e os resultados promissores dos testes preliminares. Por outro lado, no entanto, devido a limitações significativas na metodologia (particularmente a pequena quantidade de dados e a falta de validação independente), melhorias adicionais são necessárias para confirmar completamente a eficácia da solução. Na minha opinião, o trabalho requer que as correções acima sejam feitas, após o que poderá ser aceito para publicação. Recomendo complementar as análises ausentes (mais dados, possivelmente arquiteturas mais profundas, testes de validação). Uma vez implementadas, o artigo poderá ser uma contribuição valiosa para a literatura de informática médica.

**Estado final da revisão:**

O trabalho apresentado para revisão demonstra o trabalho empolgante dos autores que compuseram este manuscrito. O artigo, como trabalho de pesquisa, atende aos critérios definidos para autores serem publicados na BioMedInformatics.

**Portanto, peço ao Editor que recomende este trabalho para publicação após uma revisão importante.**

****Revisor 3****

**Caros autores**

O manuscrito intitulado “Rede neural convolucional para apoiar a tomada de decisão de enfermeiros sobre terapias tópicas para úlceras venosas: construção, validação e avaliação” aborda questões importantes relacionadas à tomada de decisão clínica assistida por IA no tratamento de feridas; no entanto, precisa de grandes melhorias com base nos seguintes comentários:

1. Como o conjunto de dados de imagens de úlceras venosas foi coletado, anotado e validado?
2. Por favor, esclareça se técnicas de aumento de dados são aplicadas para melhorar a generalização do modelo?
3. O estudo afirma ter 100% de precisão; os autores verificaram se isso é overfitting? Por favor, discuta.
4. Os autores devem esclarecer os critérios que foram usados ​​para excluir certos curativos do modelo final.
5. Por favor, esclareça como a arquitetura CNN foi especificamente adaptada para análise de imagens de feridas; quais camadas convolucionais e filtros foram usados? Como o conjunto de dados foi dividido para treinamento, validação e teste? Inclua contagens de amostra para cada subconjunto. Os autores devem adicionar as complexidades computacionais e os custos para todos os trabalhos revisados. Os autores devem adicionar os trabalhos mais recentes e relevantes para obter clareza e compreensão sobre a arquitetura CNN; divisão do conjunto de dados; e complexidade computacional [1] Tri-STANet: Uma estrutura avançada que integra transformador de eixo duplo e codificação morfológica para detecção aprimorada de AS; [2]; Métodos de Inteligência Artificial para Assistência Diagnóstica e de Tomada de Decisão em Feridas Crônicas: Uma Revisão Sistemática [3] dcnnbt: um novo modelo de classificação de tumor cerebral baseado em rede neural de convolução profunda
6. É possível atualizar o modelo com novas evidências ou tipos de curativos no futuro?
7. Que medidas foram tomadas para evitar viés nos dados de treinamento?
8. Os autores devem explicar como o aplicativo funciona em ambientes clínicos em tempo real em comparação aos métodos tradicionais.
9. Como a incerteza ou a baixa confiança nas previsões são tratadas pelo sistema?
10. Quais recursos computacionais são necessários para executar o aplicativo em ambientes clínicos típicos?
11. É necessário esclarecimento sobre a conta do modelo para fatores específicos do paciente?
12. A confiabilidade entre avaliadores foi avaliada entre os juízes especialistas durante a fase de validação?

**Revisor 4**

O artigo apresenta o desenvolvimento de uma ferramenta inteligente (VenoTEC) para auxiliar enfermeiros na seleção de terapias tópicas para úlceras venosas. A transformação de um sistema de apoio à decisão "baseado em IA" em uma **ferramenta prática, voltada para enfermeiros** , para o tratamento de úlceras venosas, com usabilidade validada.

Embora a intenção seja valiosa, o estudo sofre de sérios problemas metodológicos e de relato, pois:

1. **Limitações do conjunto de dados** – O desenvolvimento contou com apenas 3 especialistas (8 casos) e testes de usabilidade com 13 enfermeiros. Este conjunto de dados é muito pequeno e privado, sem validação externa ou pública. A precisão de 100% relatada é, portanto, irrealista e provavelmente reflete overfitting ou testes em dados de treinamento. Um conjunto de dados tão pequeno só seria aceitável se os autores também validassem sua abordagem usando um conjunto de dados maior e publicamente disponível de imagens/casos de feridas semelhantes ou relevantes. Sem isso, os resultados atuais carecem de generalização.
2. **Representação incorreta do modelo** – Apesar de ser descrita como uma CNN, a arquitetura é uma rede feedforward simples. Não são utilizadas camadas convolucionais nem extração de características de imagem.
3. **Falta de rigor no design do modelo** – A escolha da arquitetura não é adequadamente justificada. O artigo não fornece otimização sistemática de hiperparâmetros, estudo de ablação ou comparação com linhas de base simples (por exemplo, regressão logística, SVM ou CNNs baseadas em ResNet). Afirmar que "várias arquiteturas foram testadas" é insuficiente sem evidências do processo ou dos resultados. No mínimo, os autores devem demonstrar que avaliaram múltiplos tamanhos de camadas (por exemplo, adicionando/removendo neurônios e camadas) e relatar diferenças de desempenho. Além disso, como esta é uma ferramenta de suporte à decisão clínica, técnicas de IA explicável (XAI) devem ser consideradas para fornecer interpretabilidade das recomendações do modelo. Sem a transparência do modelo, enfermeiros e clínicos não podem confiar ou validar a lógica por trás dos resultados. Além disso, métodos como **aprendizagem por transferência** (usando modelos de imagens médicas pré-treinados) ou **aumento de dados baseado em GAN** poderiam ter sido explorados para superar as limitações do conjunto de dados e melhorar a robustez. Em sua forma atual, o design do modelo parece ad hoc e carece de rigor metodológico, interpretabilidade e escalabilidade.
4. **Validação fraca** – A avaliação limita-se à usabilidade (pontuação SUS) e simulações internas. Não são fornecidos testes de precisão/recall, validação cruzada ou resultados clínicos.
5. **Transparência e reprodutibilidade** – Nenhum conjunto de dados, código ou parâmetros de treinamento detalhados são compartilhados. Outro pesquisador não pode replicar os resultados.

### **Recomendação**

Recomendo **a reenvio após uma revisão importante** . Para que o artigo seja publicável, os autores devem:

* Use ou valide em um conjunto de dados maior/público.
* Descreva claramente o modelo como um MLP ou implemente uma CNN verdadeira.
* Fornece ajuste sistemático de hiperparâmetros e linhas de base comparativas.
* Inclua validação adequada (métricas, validação cruzada, resultados clínicos).
* Melhore a transparência (código, métodos, seção de Limitações explícitas).

Atualmente, o trabalho é preliminar e não está pronto para publicação