

A Model Based on Heterogeneous Data Processing for Clinical Support Applications

Rodrigo F. Rönna

Universidade do Vale do Rio
dos Sinos (UNISINOS)
São Leopoldo, Brasil
rodrigo.ronnau@gmail.com

Sandro J. Rigo

Universidade do Vale do Rio
dos Sinos (UNISINOS)
São Leopoldo, Brasil
rigo@unisinos.br

Marta R. Bez

Universidade Feevale
Novo Hamburgo, Brasil
martabez@feevale.br

Jorge L. V. Barbosa

Universidade do Vale do Rio
dos Sinos (UNISINOS)
São Leopoldo, Brasil
jbarbosa@unisinos.br

ABSTRACT

The use of information systems to aid in the practice of medicine is currently under study, aiming to evaluate its possibilities to improve the quality of patient care. Among the applications for this purpose, it is possible to highlight those that focus on the extraction of characteristics in medical texts and images. This work proposes a model that uses different input formats with the objective of providing, in an integrated way, support to the doctor or specialist.

KEYWORDS¹

Medical informatics, image processing, natural language processing

1 INTRODUÇÃO

Os sistemas médicos de informação possuem um papel fundamental no processo de atendimento aos pacientes, contribuindo na obtenção de avanços importantes de qualidade e eficiência na prestação desses serviços [1]. Podem ser encontrados estudos que tratam de aplicações baseadas em processamento de textos [2] e trabalhos fundamentados em processamento de imagens [3]. Entretanto, a prática da atividade médica demanda a integração de diferentes tipos de dados, que viabiliza a obtenção de melhorias nos resultados [4].

A primeira lacuna de pesquisa, identificada nos trabalhos relacionados [4-5], está ligada à extração e integração de informações a partir do processamento dos diferentes tipos de entradas disponíveis na área médica, como textos, imagens, vídeos, sons e sensores. Outra lacuna é a ausência de um padrão comum para a representação e o armazenamento dos dados obtidos no processamento das entradas. A terceira lacuna é a falta de um modelo que permita agregar novos recursos de maneira simples. A área clínica é bastante ampla, de modo que a possibilidade de incrementar os sistemas de apoio com o desenvolvimento de novas aplicações, que fazem uso dos dados já existentes, é um requisito desejável. Os três aspectos citados

como lacunas compõem a base de especificações do modelo proposto neste artigo.

Este trabalho apresenta, na Seção 2, o modelo proposto. Detalhes sobre o desenvolvimento do protótipo são explicados na Seção 3. Por fim, a Seção 4 contém as conclusões e sugestões de trabalhos futuros.

2 MODELO

O principal objetivo deste trabalho é apresentar um modelo que permita a utilização de diferentes tipos de entradas de dados, de maneira integrada, em aplicações de suporte clínico. Após o estudo da bibliografia foi possível determinar os elementos que o modelo deve conter, bem como o relacionamento entre eles, de modo que seja viável alcançar o objetivo citado. A Figura 1 apresenta a estrutura definida.

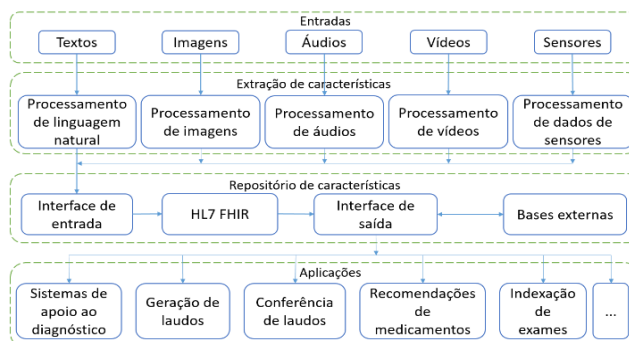


Figura 1: Estrutura do modelo proposto.

Os elementos que compõem o grupo denominado Entradas representam os diferentes formatos de informações que podem ser utilizados na obtenção de características. Cada um destes tipos é processado por uma rotina correspondente, que tem como finalidade extrair as informações que poderão ser utilizadas pelas aplicações. Esses processos estão agrupados no bloco denominado Extração de características. Essas rotinas têm, como saída, as características que serão armazenadas no repositório, de maneira estruturada. A possibilidade de empregar qualquer tipo de dado disponível na área da saúde visa atender a primeira lacuna citada, que trata da utilização de fontes heterogêneas.

Com o objetivo de suprir a segunda lacuna de pesquisa, que trata da interoperabilidade entre sistemas, e de armazenar as características extraídas de maneira estruturada, é utilizado o

In: Sessão de Pôsteres do WebMedia'17, Gramado, Brasil. Anais do XXIII Simpósio Brasileiro de Sistemas Multimídia e Web: Workshops e Pôsteres. Porto Alegre: Sociedade Brasileira de Computação, 2017.

padrão HL7 FHIR. As interfaces de entrada e saída possibilitam a compatibilização entre o HL7 FHIR e as aplicações que foram desenvolvidas empregando outros padrões. A possibilidade de integração com bases externas é representada pelo componente de mesmo nome, viabilizando a utilização de informações disponíveis em outros repositórios.

O último grupo cita exemplos de aplicações que fazem uso dos dados armazenados no repositório de características, como os sistemas de apoio ao diagnóstico, conferência de laudos, recomendação de medicamentos, entre outros. A partir dessas funcionalidades, que possuem diferentes finalidades, é demonstrado o preenchimento da terceira lacuna de pesquisa, que trata da possibilidade de estender o emprego do sistema.

3 PROTÓTIPO

Com a finalidade de verificar a capacidade de utilização e a integração dos itens que formam o modelo, foi desenvolvido um protótipo que permite detectar, em laudos de exames por imagem, se informações importantes não foram citadas e se dados incorretos foram inseridos. A Figura 2 apresenta os componentes do modelo, destacados em vermelho, que compõem o protótipo construído.

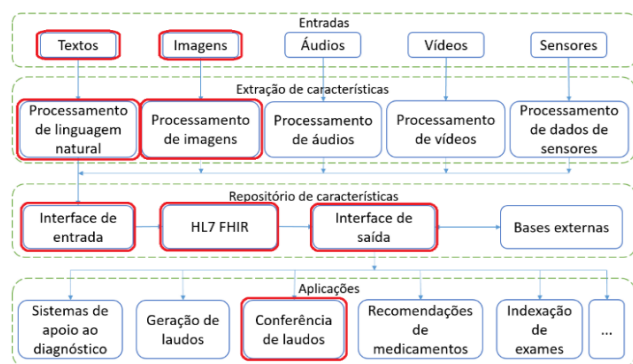


Figura 2: Elementos desenvolvidos no protótipo.

O primeiro componente desenvolvido foi o repositório de características. Foi utilizada a biblioteca HAPI FHIR² (implementação Java da especificação HL7 FHIR) nas interfaces de entrada, saída e no servidor que armazena as características.

A segunda etapa envolveu a implementação relacionada ao processamento de imagens de tomografia computadorizada, com base na ferramenta descrita por Reis et al. [6]. Foi incluído nesse sistema o processo de identificação de pontos de calcificação a partir do método descrito por Flores [7].

A etapa seguinte consiste no processamento do laudo correspondente às imagens avaliadas. O texto é submetido ao PALAVRAS [8], que disponibiliza as palavras classificadas de acordo com aspectos sintáticos e morfológicos. Os substantivos das frases que contêm o termo “calci” são verificados através do UMLS *Metathesaurus Browser*, buscando identificar se estão

relacionados às regiões anatômicas de interesse. Se o trecho não estiver sendo negado, a referência à presença de calcificação é considerada válida e é adicionada no repositório de características.

O último elemento criado foi a aplicação de conferência de laudos. Para cada exame são feitas consultas ao repositório das características, possibilitando comparar os resultados obtidos no processamento de cada tipo de entrada (imagem e texto). Se forem detectadas diferenças é exibido um alerta para o usuário responsável, indicando a situação identificada e sugerindo que o exame seja revisado.

4 CONCLUSÃO

Neste artigo foi apresentado um modelo em desenvolvimento baseado no processamento de dados heterogêneos para aplicações de suporte clínico. O protótipo construído é composto por elementos existentes em todas as camadas do modelo proposto, demonstrando sua funcionalidade e evidenciando a sua importância. Além disso, cada componente do sistema foi desenvolvido de forma independente, com base em aplicações já validadas e apresentadas na literatura. Isso demonstra a característica modular e expansível da estrutura proposta. O uso de diferentes tipos de dados como entrada e do padrão HL7 FHIR preenchem, respectivamente, as outras duas lacunas de pesquisa descritas.

Atividades futuras envolvem a avaliação do modelo, o desenvolvimento de outras aplicações, experimentos adicionais para confirmar a sua utilidade, viabilidade, facilidade de uso e limitações, bem como o emprego em um cenário real de uso.

AGRADECIMENTOS

Agradecemos à CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pelo aporte financeiro oferecido para a realização desta pesquisa.

REFERÊNCIAS

- [1] R. G. Fichman, R. Kohli and R. Krishnan. 2011. Editorial overview - the role of information systems in healthcare: current research and future trends. *Information Systems Research* 22, 3 (2011), 419–428.
- [2] E. Pons et al. 2016. Natural language processing in radiology: a systematic review. *Radiology* 279, 2 (2016), 329–343.
- [3] M. Firmino et al. 2016. Computer-aided detection (CADe) and diagnosis (CADx) system for lung cancer with likelihood of malignancy. *Biomedical engineering online* 15, 1 (2016).
- [4] M. Puppala et al. 2015. METEOR: an enterprise health informatics environment to support evidence-based medicine. *IEEE Transactions on Biomedical Engineering* 62, 12 (2015), 2776–2786.
- [5] M. Williams et al. 2009. A modular framework for clinical decision support systems: medical device plug-and-play is critical. *ACM SIGBED Review* 6, 2 (2009), 8.
- [6] A. N. Reis et al. 2016. Ferramenta para apoio ao diagnóstico baseada em processamento de imagens de tomografia computadorizada do tórax. In *IV ERCAS - Escola Regional de Computação Aplicada à Saúde*, 20-21.
- [7] J. C. Flores. 2015. Detecção Automática de Calcificação Arterial Coronariana em Tomografia Computadorizada. Trabalho de Conclusão de Curso (Monografia) – Curso de Ciência da Computação, Universidade Feevale, Novo Hamburgo, RS.
- [8] E. Bick. 2000. The Parsing System “PALAVRAS”: Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework. Aarhus Universitetsforlag.

² HAPI FHIR – <http://hapifhir.io/>