

Universidade de São Paulo
Plano de Gestão de Dados

Identificador: 123456

ORCID:XXXX-XXXX-XXXX-XXXX

Pesquisador responsável: José da Silva Sauro

Contato para responsável pelos dados: joaosauro@usp.br

Data de criação do plano: xx/xx/xx

Local: Laboratório de Análise de Imagens, Escola de Pesquisa Científica Fictícia,
Universidade de São Paulo

Título do Projeto: Geração sintética de imagens para avaliação de sistemas de auxílio
ao diagnóstico

Descrição do Projeto: este projeto visa a desenvolver algoritmos que, a partir de
parâmetros fornecidos pelo usuário, gera imagens sintéticas para simular radiografias
pulmonares com achados relacionados a anomalias.

Financiador(es): Agência XYZ

1) Dados gerados

Os dados gerados pelo projeto em questão serão de dois tipos:

- a) Imagens médicas sintéticas, criadas a partir sistema computacional com a finalidade de simular casos de diagnóstico de câncer de pulmão. Essas imagens serão geradas no formato JPEG sem perda, com resolução de contraste de 16 bits e resolução espacial de 1200 dpi. Serão geradas duas mil imagens no formato citado, disponibilizadas em um único arquivo no formato .zip sem perda. Cada imagem será identificada por um número único.
- b) Imagens reais digitalizadas a partir de filme radiográficos disponibilizados pelo Hospital XYZ. Essas imagens serão disponibilizadas no formato JPEG sem perda, com resolução de contraste de 16 bits e resolução espacial de 1200 dpi. Serão disponibilizadas 500 imagens no formato citado, em um único arquivo no formato .zip sem perda. Cada imagem será identificada por um número único.
- c) Dados de diagnóstico relacionados às imagens disponibilizadas. Esses dados serão disponibilizados em formato csv, com as seguintes colunas:

- a. número da imagem: número inteiro que identifica a imagem no arquivo disponibilizado nos itens (a) e (b);
- b. coordenadas iniciais da localização do achado (no formato XXXX,YYYYY: em que XXXX indica a linha da imagem e YYYYY indica a coluna da imagem);
- c. coordenadas finais da localização do achado (no formato XXXX,YYYYY: em que XXXX indica a linha da imagem e YYYYY indica a coluna da imagem)
- d. descrição do achado: texto com no máximo 200 caracteres que descreve o achado, extraído do laudo médico no caso de imagens do item (b) ou a simulação realizada no caso de imagens do grupo (a).

2) Localização dos dados

Os dados serão disponibilizados no repositório de dados científicos da Universidade de São Paulo (<http://dadoscientificos.usp.br>).

As imagens sintéticas (item 1a) não envolvem aspectos éticos ou legais. Podem ser utilizadas em projetos de pesquisa, desde que a fonte (projeto presente) seja citada. Podem igualmente sempre reproduzidas em documentos públicos, tais como artigos científicos e relatórios técnicas, desde que citada a fonte.

As imagens reais (item 1b) estão devidamente anonimizadas e sua aquisição foi aprovada pelo Comitê de Ética do Hospital XYZ. Podem ser utilizadas para processamento em projetos de pesquisa, desde que a fonte (projeto presente) seja citada. No entanto, devido a aspectos legais essas imagens não podem ser reproduzidas em documentos públicos, tais como artigos e relatórios científicos.