



Universidade do Minho
Departamento de Informática
Mestrado [Integrado] em Engenharia Informática

Perfil SISTEMAS INTELIGENTES
Aprendizagem Profunda (Deep Learning)
1º Ano, 2º Semestre
Ano letivo 2024/2025

Trabalho de Grupo – Modulo 2
Abril, 2025

Tema

Modelação de Técnicas Cirúrgicas em Vídeo via Deep Learning - Conceção e otimização de uma solução de visão por computador para análise e interpretação de vídeos de treino em tarefas de sutura.

Contexto do Desafio

O desafio OSS: Open Suturing Skills (parte da Endoscopic Vision Challenge 2025) visa desenvolver sistemas de visão computacional para análise automática de vídeos de treino de sutura cirúrgica. O objetivo é fornecer feedback quantitativo sobre a qualidade técnica do procedimento, identificando:

- Fases da sutura (e.g., entrada da agulha, tração do fio)
- Erros técnicos (e.g., ângulo incorreto da agulha, tensão inadequada do fio)
- Métricas de desempenho (e.g., tempo por fase, fluidez do movimento)

Objetivos de Aprendizagem

Este trabalho prático visa sensibilizar e motivar os alunos para a conceção e desenvolvimento de soluções baseadas em Aprendizagem Profunda, aplicando técnicas abordadas ao longo do semestre, bem como outras descritas na literatura. O desafio consiste em conceber um modelo/aplicação/protótipo com potencial para participação no subdesafio "OSS: Open Suturing Skills" do Endoscopic Vision Challenge a realizar-se na MICCAI 2025.

Especificamente tem-se como objetivos:

- Aplicar técnicas de Deep Learning (CNNs, Transformers, LSTMs, VIT, vLMs) em vídeos médicos
- Dominar fluxos de pré-processamento para dados multimodais (RGB, infravermelho, metadados temporais)
- Implementar métricas específicas para avaliação de competências cirúrgicas
- Garantir reprodutibilidade através de pipelines bem documentadas
- Preparar submissões técnicas seguindo padrões de competições internacionais

Enunciado

Propõe-se a conceção/desenvolvimento de uma solução computacional, com base em técnicas de Deep Learning, para participação no subdesafio "OSS: Open Suturing Skills" incluído na "Structured description of the challenge design of the Endoscopic Vision Challenge", que terá lugar durante a conferência MICCAI 2025, em Daejeon, República da Coreia.

As tarefas a planear/desenvolver poderão incluir:

- Exploração, análise e preparação do dataset disponibilizado pelo desafio, com vista à extração de conhecimento relevante;

- Conceção e otimização de um modelo de Aprendizagem Profunda para o problema proposto, incluindo, se pertinente, modelos tradicionais de machine learning como baseline;
- Avaliação comparativa dos modelos desenvolvidos, utilizando métricas apropriadas e técnicas robustas de estimativa de erro;
- Interpretação dos resultados obtidos e da performance do(s) modelo(s) à luz da análise exploratória inicial;
- Opcionalmente, desenvolvimento de uma aplicação (desktop ou web) que permita a utilização do modelo final por utilizadores finais (frontend à escolha do grupo).

Enunciado Técnico

Conceber/desenvolver um sistema de IA para análise automática de vídeos de sutura, capaz de:

Tarefas (a acertar após a divulgação pública do *challenge* no site):

1. Segmentação Temporal
 - Identificar fases da sutura no vídeo
 - Classificar transições entre etapas usando modelos de série temporal
2. Análise Qualitativa
 - Detetar erros técnicos com base em padrões cinemáticos
 - Quantificar métricas de desempenho (e.g., precisão angular, velocidade)
3. Interpretabilidade
 - Gerar mapas de saliência destacando regiões críticas no vídeo
 - Produzir relatórios automáticos para orientação do formando

Metodologia

A edição de 2025 do subdesafio "OSS: Open Suturing Skills" será idêntica à da edição de 2024. Pelo que, inicialmente, devam ler atentamente todas as páginas respetivas, nomeadamente "How to Participate", "Tasks", "Data" e "Evaluation" em <https://www.synapse.org/Synapse:syn54123724>:

1. "How to Participate" – Quando a página da edição de 2025 estiver disponível, essa informação será disponibilizada e um dos elementos do grupo deve prontamente registar-se e fazer o pedido para aceder ao dataset. Recomenda-se que os restantes elementos do grupo também se registem para terem acesso a todas as funcionalidades.
2. "Tasks" – Nesta página é apresentada a descrição das tarefas a realizar. Será explicado o objetivo de cada tarefa, o qual deve ser utilizada como foco na vossa pesquisa.
3. "Data" – Nesta página é apresentada a descrição do dataset. O melhor método para processamento de dados é muito dependente do tipo de dados que se esteja a utilizar, e.g., volumes, vídeos, imagens, tabelares, etc. Devem focar a pesquisa na forma mais eficiente de explorar o tipo de dados fornecidos (neste caso vídeos e tabelares) e a melhor forma de processar esses dados.
4. "Evaluation" – Nesta página é apresentada a descrição das métricas que serão utilizadas para avaliar os resultados. Estas métricas devem ser implementadas por cada grupo de forma a testar as soluções.

É importante que todos os elementos do grupo leiam com atenção as páginas mencionadas. Visto que, por enquanto, não existe acesso aos dados, podem utilizar os dados disponíveis em <https://zenodo.org/records/7940583> para se ambientarem ao tipo de dados que serão fornecidos (basta fazer o download de um dos "Package" disponíveis). A pesquisa de possíveis soluções deve ser orientada ao objetivo do desafio e ao tipo de dados com que terão de trabalhar.

Pelo que o workflow sugerido é:

1. Leitura das páginas mencionadas;
2. Revisão da literatura:
 - a. da melhor forma de trabalhar com o tipo de dados fornecidos;

- b. de soluções existentes que podem ser aplicadas diretamente ao problema;
 - c. de soluções que podem ser adaptadas ao problema.
3. Implementação da solução:
 - a. Criação de notebooks para pre-processamento dos dados, desenvolvimento da solução e avaliação da solução.
4. Publicação do repositório e escrita do relatório técnico.

De notar que é recomendável iniciar a escrita do relatório pouco depois de iniciar a revisão da literatura. O código deve ser bem estruturado, organizado e guardado no repositório público de forma frequente para evitar perdas e confusões. Por norma, as soluções demoram vários dias a ser desenvolvidas, por exemplo soluções baseadas em aprendizagem profunda que demoram vários dias a treinar. Visto isto, o desenvolvimento do projeto deve ser programado para dedicar tempo suficiente a estas tarefas.

Entrega e Avaliação

Os resultados do trabalho deverão ser apresentados num relatório técnico, com um máximo de 10 páginas, contendo:

- Metodologia seguida para a análise dos dados e desenvolvimento do software;
- Descrição geral do software/scripts desenvolvidos (detalhes adicionais devem constar em documentação própria no repositório);
- Análise exploratória dos dados utilizados e dos tratamentos aplicados;
- Detalhes sobre os modelos desenvolvidos: arquitetura, tuning de hiperparâmetros (se aplicável), gráficos do processo de treino, entre outros;
- Sumário dos resultados obtidos e respetiva análise crítica;
- Recomendações e sugestões com base nos resultados.

Requisitos adicionais:

- Todo o código desenvolvido deve ser disponibilizado num repositório público (por exemplo, GitHub), acompanhado da documentação e dados necessários para garantir a reprodutibilidade dos resultados.
- Para as análises de dados, recomenda-se o uso de Jupyter Notebooks ou ferramentas equivalentes, de modo a facilitar a transparência e replicação.

Nota Pedagógica

Este enunciado segue a filosofia challenge-based learning, onde os alunos não apenas desenvolvem soluções, mas também aprendem a:

- Trabalhar com dados desbalanceados e anotações esparsas
- Implementar fairness checks para evitar viés em avaliações de competência
- Escrever relatórios técnicos no formato MICCAI (template LaTeX disponível no site <https://conferences.miccai.org/2025/en/PAPER-SUBMISSION-GUIDELINES.html#manuscriptpreparation>)

Para dúvidas técnicas, os alunos podem aceder ao fórum do desafio que será disponibilizado na página do desafio, em "Discussion".

No dia **29 de maio de 2025** decorrerão as sessões de apresentação do trabalho desenvolvido. Os grupos de trabalho deverão submeter antecipadamente as suas apresentações na plataforma de e-learning da Universidade do Minho. Deste modo evita-se andar a trocar de computador no dia da apresentação. Cada grupo disporá de 10 minutos para realizar a apresentação, utilizando os meios que considerar mais adequados, seguindo-se 5 minutos para discussão.

O relatório deverá ser submetido, por um elemento do grupo, na plataforma de *e-learning* da Universidade do Minho, indicando nessa submissão o URL do repositório do software desenvolvido. As submissões deverão ser realizadas até ao final do dia **2 de junho de 2025**. Os repositórios poderão também ser atualizados até essa data. Dá-se assim possibilidade de ainda terem em conta o feedback recebido na apresentação.

Referencias

<https://opencas.dkfz.de/endovis/challenges/2025/>
https://x.com/endo_vis
<https://miccai.org/index.php/special-interest-groups/challenges/miccai-registered-challenges/>
<https://conferences.miccai.org/2025/en/default.asp>
<https://opencas.dkfz.de/endovis/challenges/2024/>
<https://zenodo.org/records/15075458>
<https://conferences.miccai.org/2025/en/challenges.asp>
<https://chatpaper.com/chatpaper/pt/paper/116965>
https://openreview.net/group?id=MICCAI.org%2F2025%2FWorkshop%2FMSB_EMERGE

Código de Conduta

Os participantes do presente trabalho académico declaram ter atuado com integridade e confirmam que não recorreram à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração. Mais declaram que conhecem e respeitaram o Código de Conduta Ética da Universidade do Minho.

Avaliação pelos pares

Cada grupo deverá realizar uma análise coletiva sobre o contributo e esforço que cada elemento deu para o avanço do trabalho. Dessa análise devem conseguir identificar os membros que trabalharam acima, na e abaixo da média. Para esta componente de avaliação está previsto 1 valor para cada aluno (5% da avaliação) que reflete a sua contribuição individual no desenvolvimento do instrumento.

Para tal, cada grupo deverá enviar um email para valves@di.uminho.pt até ao dia **2 de junho de 2025**, com todos os elementos da equipa em CC. O assunto deverá ser "AP - Avaliação pares".

No texto do email deverão indicar para cada elemento do grupo, o respetivo delta (parcela a somar à nota desta componente). Lembra-se que os delta podem ser negativos, nulos ou positivos e que, em cada grupo, o somatório dos deltas deve ser igual a 0.00

Exemplo 1 (todos os alunos recebem 1 valor correspondendo a um esforço igual entre todos):

PG1234 João Silva DELTA=0
PG5678 António DELTA=0
PG9123 Maria DELTA=0

Exemplo 2 (O António recebe 2 valores, o João e a Maria recebem 0.5 valores nesta componente):

PG1234 João Silva DELTA= -0.5
PG5678 António DELTA= 1
PG9123 Maria DELTA= -0.5