Work package description

- **Lead beneficiary**: OBU (Oxford Brookes University)
- Prof Fabio Cuzzolin, WP Leader
- AI and Vision group, School of Engineering, Computing and Mathematics, Oxford Brookes University
- Phone: +44 (0)1865 484526
- Email: fabio.cuzzolin@brookes.ac.uk

- Partners involved and their role (in **Red M1-M12**)

<table>
<thead>
<tr>
<th>Partner</th>
<th>PM</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNIVR</td>
<td>4</td>
<td><strong>Current stage recognition</strong>, predicting future surgeon action, decision making an intervention, anomaly detection.</td>
</tr>
<tr>
<td>UNIFE</td>
<td>2</td>
<td><strong>Current stage recognition</strong>, predicting future surgeon action, decision making an intervention, anomaly detection.</td>
</tr>
<tr>
<td>OSR</td>
<td>3</td>
<td><strong>Online surgeon action recognition, current stage recognition</strong>, predicting future surgeon action, decision making an intervention, anomaly detection.</td>
</tr>
<tr>
<td>FSCR</td>
<td>1</td>
<td>Predicting future surgeon action, decision making an intervention, anomaly detection.</td>
</tr>
<tr>
<td>UPC</td>
<td>6</td>
<td>Predicting future surgeon action, decision making an intervention, anomaly detection.</td>
</tr>
<tr>
<td>OBU</td>
<td>42</td>
<td><strong>Online surgeon action recognition, current stage recognition</strong>, predicting future surgeon action, decision making an intervention, anomaly detection.</td>
</tr>
</tbody>
</table>
Milestones:

- **MS4**: the *MULTIROBOTS-SURGERY* platform is operative at Month 12. Experimental/clinical validations are conducted by OSR surgeons at UNIVR.

Deliverables:

- **D6.1 Real-time surgeon action detection and recognition** (Month18). Development of tools for surgeon action detection and recognition based on a novel deep learning architecture able to regress ‘action tubes’ in real time from incoming videos.
Title: **Online surgeon action recognition** [M1-M18]

Participants: OBU, OSR

- **Activities:**
  - The real-time detection (in time), localisation (within the image) and classification of multiple actions and events
  - Actions are performed either by the main surgeon or by the assistive robotic arms
  - Actions can be: dissecting a tissue line, sucking up blood, manipulating anatomical structures, changing camera view, etcetera
  - Events can be: excessive bleeding, accidental cuts, etc

- **Challenges:**
  - We do have a real-time pipeline already in place: **what we need is domain-specific data the system can learn from (data scarcity)**
  - Resolution/quality of the endoscope images
  - To detect anomalous events we need real data
Task 6.1 - Method

- Works by detecting actions of interests in each frame, in real time
- Then linking them together, to form coherent ‘action tubes’
Task 6.1 - Illustration

• The system learns from ‘annotated’ videos: somebody draws a bounding box around actions of interest
• And ‘labels’ them (e.g. this is jumping, this is dancing)
• We need something similar for surgical data
• The more/better annotation we have, the better the results
• May come from verbal annotation during the procedure
• Plus a simple graphical interface to draw the boxes
• A doctoral student from OSR could do this?
Task 6.2

Title: **Current procedure’s stage recognition [M1-M24]**
Participants: OBU, OSR, UNIFE, UNIVR

- **Activities:**
  - Based on prior knowledge on the temporal structure of a laparoscopic procedure, and on (training) examples of similar procedures..
  - .. the AI core recognises, at any time instant, what stage of the procedure it is in
  - Arranges individual actions (see Task 6.1) into a graph encoding the spatial and temporal structure of the procedure
  - Makes use of both examples (as in Task 6.1) and of **prior clinical knowledge** (experts describing the structure of a procedure)

- **Challenges:**
  - **Scene context is needed** (coming from WP5), config. of surgical cavity
  - **Not implemented yet** (but we have a clear architecture in mind)
  - **Real time can be challenging** (but lots of prior knowledge)
- **Nephrectomy example**
- The same intervention can happen in different variants
- Stages are identified based on what is happening (T6.1), tools present (T5.3) - need to associate visual info with procedure stages
- Prior clinical knowledge on possible temporal structures is crucial
WP relationship (M1-M12)

- **Surgical procedure specification and validation (OSR)**
- **Anatomical modelling, phantom (UNIVDUN)**
- **Multi-modal human-robot interface and bilateral teleoperation (UNIMORE)**
- **Cognitive control, task supervision and computer vision (UNIVR)**
  - System integration (UNIVR)

**WP1**
- Surgical knowledge on temporal structure of procedures, possible variants, etc
- Finite state machine models (M6)
- Any training data coming from real procedures?

**WP2**
- Multirobot surgery platform (M12)
- Training data generated by using the platform on phantoms

**WP3**
- Detection of tools, scene understanding (M24), for stage recognition
- Tracking of tools and tissue (M30), for anomaly detection

**WP4**
- Anomalies detected (M18)
- Decisions on action to perform in SOLO-SURGERY scenario (M24)
- Decisions on action to perform in LAPARO2.0-SURGERY scenario (M24)

**WP5**
- Technical system specifications (M6)
- Software/Hardware Architecture (M12)

**WP6**

**WP7**
- Multi-robot Cooperation and Task Planning (UPC)
Facilities @ AI and Vision lab

- **Already available:**
  - 1-GPU workstation, 4.2TB HD (Jupiter)
  - Dell Precision Tower 7910 Workstation equipped with two NVIDIA TITAN X GPU cards, 12.5 TB HD (Sun)
  - OcUK Tech Labs Haswell-E X99 Pro Gaming PC Configurator with 4 GeForce GTX 1080Ti "Blower Edition" GPUs, 19.5TB HD (Mars)
  - Two Latitude laptops for visitors
  - Annotated datasets: UCF-101, JHMDB-51, Oxford RobotCar

- **Funded by SARAS:**
  - A new 4/8 GPU workstation for deep learning processing
Risks and criticalities

- **Major risks**
  - Data of insufficient quantity and/or quality is collected to train the deep networks performing Tasks 6.1 and 6.2 - > **MULTIROBOTS-SURGERY needed asap, appropriate data collection protocol to be agreed on, clinical support by OSR UNIVR**
  - Inaccurate or slow recognition of procedure stage - > The workflow of the R-MIS procedure is much constrained by clinical knowledge, so easing the design of complex procedure models. Efficiency by fast deep neural nets
  - Anomalies go undetected - > **the system must be tilted towards preventing false negatives**, at least initially

- **Specific needs**
  - Clinical knowledge encoded into **temporal structure of procedures and variants**
  - **Training data from example procedures**: endoscope videos, robot signals, from the MULTIROBOTS-SURGERY platform
  - **Training data from real laparoscopies** conducted at OSR or elsewhere
  - For each example procedure: **annotation (what is happening, what objects are present, and where, what stage are we in)** – part of this we can do (object detection), most needs to come from clinical experts (verbal annotation while operating?)