**Appendix B: Ethics, Dataset Licensing, and Compliance**

This appendix outlines the ethical considerations, licensing details, and institutional compliance measures related to the datasets and experimental procedures used throughout this research.

**B.1 Dataset Licensing and Access**

This study employed two publicly available and academically recognized ultrasound image datasets:

📂 Breast Ultrasound Images (BUI) Dataset

* Source: The dataset titled “Breast Ultrasound Images (BUI)” is publicly available via [Kaggle](https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset).
* Compiled by: The dataset was compiled by Walaa Al-Dhabyani, Mohamed Gomaa, Heba Khaled, and Aly Fahmy, researchers affiliated with Egypt’s Assiut University.
* License: The dataset is released under the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors are properly credited.
* Access Link: <https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset>
* Description: This dataset includes 780 breast ultrasound images collected from 600 female patients aged between 25 and 75 years. The images are categorized into three classes: normal, benign, and malignant. Each image is accompanied by a corresponding binary ground truth mask that delineates the lesion area. The dataset is widely used for machine learning tasks including classification, lesion detection, and segmentation in breast ultrasound imaging.
* Citation: Al-Dhabyani, W., Gomaa, M., Khaled, H. and Fahmy, A., 2020. *Dataset of breast ultrasound images*. Data in Brief, 28, p.104863. Available at: <https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset> [Accessed 21 May 2025].

📂 Polycystic Ovary Syndrome (PCOS) Ultrasound Dataset

* Source: The dataset titled *"PCOS detection using ultrasound images"* is hosted on [Kaggle](https://www.kaggle.com/datasets/anaghachoudhari/pcos-detection-using-ultrasound-images).
* Compiled by: The dataset was uploaded by Kaggle user [Anagha Choudhari](https://www.kaggle.com/anaghachoudhari).
* License: The dataset is available under Kaggle's standard terms of use. While no specific license is mentioned on the dataset page, it is intended for academic and non-commercial research purposes.
* Access Link: <https://www.kaggle.com/datasets/anaghachoudhari/pcos-detection-using-ultrasound-images>
* Description: This dataset comprises ultrasound images categorized into two classes: 'infected' (indicative of PCOS) and 'notinfected' (normal). It is suitable for developing and evaluating machine learning models for PCOS detection.
* Citation: Choudhari, A., *n.d.* PCOS detection using ultrasound images [dataset]. Kaggle. Available at: <https://www.kaggle.com/datasets/anaghachoudhari/pcos-detection-using-ultrasound-images> [Accessed 21 May 2025].

Both datasets are distributed under terms that explicitly allow use in academic, non-commercial research. No license fees, IRB approvals, or clinical partnerships were required for usage under the declared conditions.

**B.2 Ethical Use of Clinical Data**

All datasets used in this research consist of fully anonymized ultrasound images that are publicly available and intended for academic use. No personal identifiers, patient metadata, or clinical records are included in either the BUI or PCOS datasets. As such, no additional institutional ethics approval or participant consent was required for their reuse in this study.

Ethical use of these datasets aligns with accepted academic standards and non-human subject research principles. Specifically:

* The datasets were anonymized by their original contributors prior to public release.
* No attempt was made to re-identify patients or reconstruct clinical histories.
* The research relied solely on open-access images provided under terms allowing non-commercial, academic research use.

In line with these safeguards, the study meets the definition of non-human subject research under international research ethics norms and common academic integrity expectations.

**B.3 Institutional Compliance**

This research adheres to the University of Canberra’s guidelines for responsible conduct of research and thesis data management. Key compliance practices include:

* Public, anonymized datasets were used exclusively, with appropriate citations and respect for stated licensing conditions.
* No modification or re-identification of sensitive clinical images was performed at any stage.
* All MATLAB scripts and experimental data were stored securely in the candidate’s university-provided directories (e.g., staff files and OneDrive), ensuring data integrity and protection.
* Each dataset used was properly attributed according to academic citation standards, with licensing information confirmed via its official distribution page (Kaggle).

These practices reflect a responsible and transparent research process appropriate for higher degree research work involving publicly available biomedical data.