City University of London

Department of Computer Science

Unsupervised Anomaly Detection for Medical Imaging

2024–2025

Student: Anndischeh Mostafavi – 220051723
Supervisor: Dr. Atif Riaz

Typeset using LaTeX

1 Introduction

This study aims to develop and evaluate an unsupervised anomaly detection system for brain tumors using *Generative Adversarial Networks (GANs)* [1]. The `GAN` will be trained on the **BraTS-2021** dataset [2], learning the distribution of healthy brain structures in *multi-modal MRI* scans to identify regions deviating from normality—potentially indicating tumor presence—without relying on labeled data (see Figure 1).

The primary outcome will be a functional GAN-based anomaly detection model that produces anomaly maps to visually highlight regions of deviation. The system is intended to assist radiologists by enhancing subtle anomaly detection. Ultimately, this research could benefit patients via earlier, more accurate diagnosis and contribute to the medical imaging community through public dissemination of the results.

The project scope includes data preprocessing (e.g., intensity normalization), selection of an appropriate `GAN` architecture such as `DCGAN` [3], `cGAN` [4], or more advanced variants (as illustrated in Figure 2), model training, and implementation of anomaly scoring mechanisms (e.g., reconstruction error or discriminator output). Evaluation will be based on established metrics such as `ROC-AUC`, precision, recall, and possibly Dice coefficient if segmentation is implemented. This work is limited to `BraTS` multi-modal MRI data and does not include clinical validation.

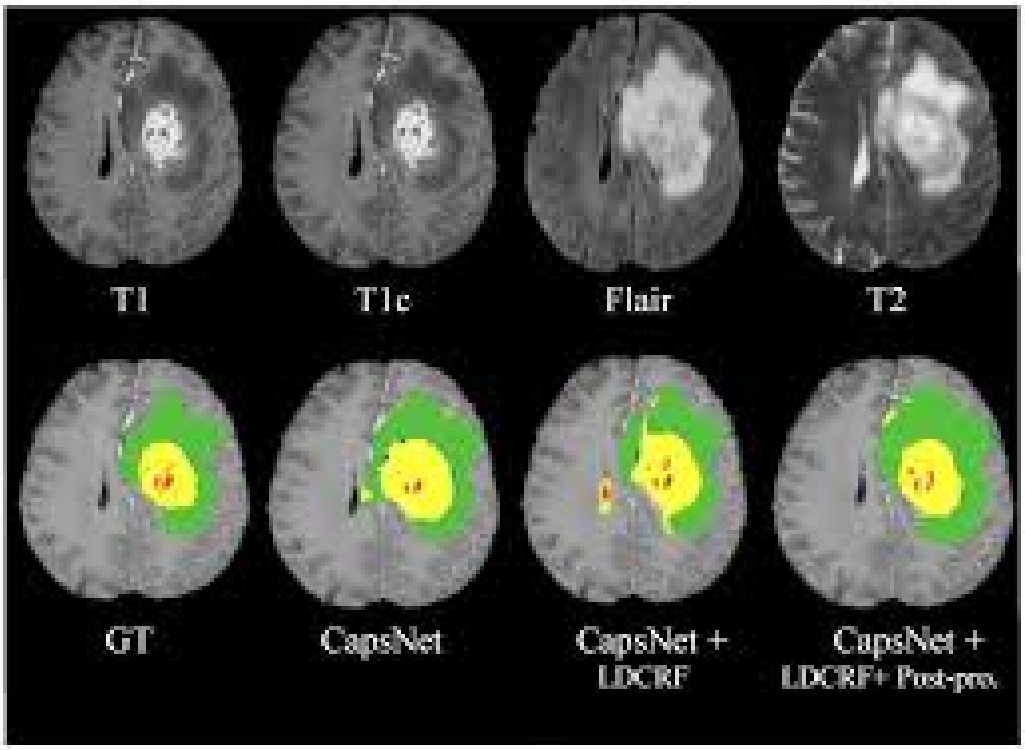

Figure 1: RSNA-ASNR-MICCAI Brain Tumor Segmentation (BraTS)[5].

2 Critical Context

The context is driven by the inherent limitations of supervised methods, which rely on large, manually annotated datasets, a costly and time-consuming process, particularly in medical imaging where expert annotation is required. This research directly addresses this challenge by exploring

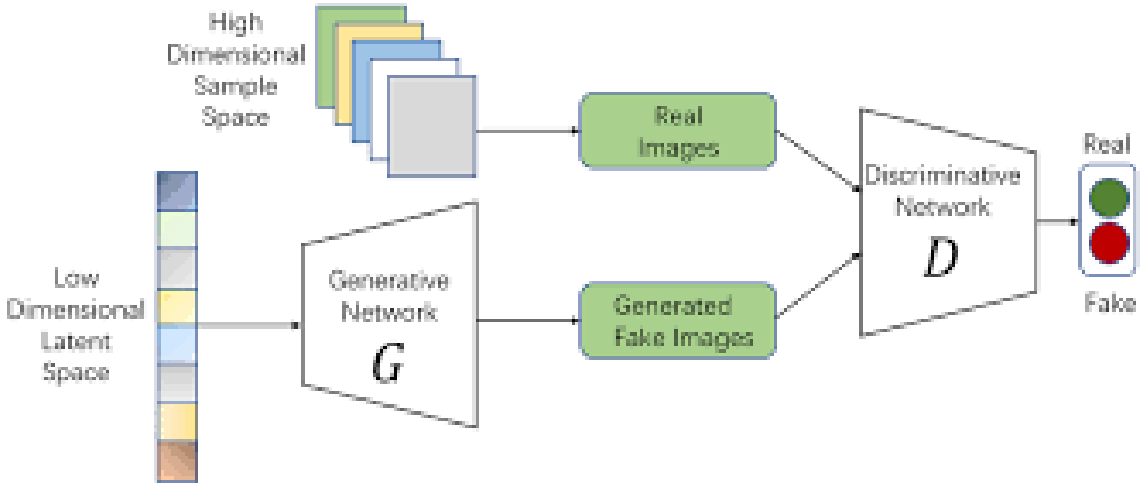

Figure 2: Architecture of Generative Adversarial Networks (GANs)[6].

the potential of GANs to learn the underlying distribution of normal brain tissue and to identify deviations indicative of tumors in an unsupervised manner.

Several recent studies have investigated the use of deep learning, including GANs, for medical image analysis. For instance, Bhandari et al. (2020) demonstrated the effectiveness of deep *convolutional neural networks (CNNs)* for brain tumor segmentation [7], highlighting the advancements of deep learning in this area. Further, Goodfellow et al. (2014) introduced the fundamental concept of GANs, establishing the basis for their application across various domains, including medical imaging [1]. Building on this, several researchers have adapted the *GAN* framework for anomaly detection. Schlegl et al. (2017) introduced *AnoGAN*, a GAN-based approach that leverages the reconstruction error from a trained generator to identify anomalous regions in medical images, demonstrating its potential for identifying abnormalities [8]. Similarly, Shobhita et al. (2020) proposed a GAN-based method for lesion detection in chest X-rays, further showcasing the versatility of GANs in medical applications [9]. Furthermore, studies have explored the use of conditional GANs (cGANs), incorporating additional information such as image modalities or contextual features to improve the performance of anomaly detection [4].

This proposal will extend the previous work by employing and comparing different GAN architectures and anomaly scoring mechanisms, within the context of the BraTS dataset. TheBraTS dataset, provides a valuable benchmark for this type of research. The utilization of this data allows the validation of the performance of anomaly detection methods, and, ultimately, the proposed approach. Finally, the work must consider the inherent limitations of the BraTS dataset in providing representative examples of normal brain tissue for all possible image variations, and tumor characteristics in order to ensure that an accurate model is built.

3 Approaches: Methods & Tools for Design, Analysis & Evaluation

This section details the technical methods and tools to be employed to achieve the project objectives. The proposed approach focuses on adapting and applying the GANomaly architecture [10], a prominent GAN-based method for unsupervised anomaly detection reviewed in [11], to the BraTS medical imaging dataset. A small novelty will be explored by comparatively evaluating different anomaly scoring mechanisms derived from the trained GANomaly model components, building upon the scoring methods discussed in the survey [11].

3.1 Data Acquisition and Preprocessing

Acquisition: Access to the `BraTS` dataset (specify the version, e.g., `BraTS 2021`) will be obtained following the challenge guidelines and data use agreements. The dataset typically includes multi-modal MRI scans (T1, T1ce, T2, FLAIR) and expert annotations (ground truth masks) for different brain tumor sub-regions.

Preprocessing: Raw `BraTS` data requires significant preprocessing to be suitable for deep learning models. This will involve: Co-registration: Aligning all four modalities for each patient volume to a common anatomical space (e.g., the T1 image).

Resampling: Standardizing `voxel` spacing and image dimensions to a uniform resolution (e.g., 1mm isotropic) and shape to ensure consistency across the dataset. Skull Stripping: Removing non-brain tissue to focus the analysis solely on the brain region.

Intensity Normalization: Standardizing intensity values across different scans and modalities (e.g., Z-score normalization or histogram matching) to mitigate scanner-specific variations.

Data Partitioning (Train/Test): The core challenge for unsupervised anomaly detection on `BraTS` is defining 'normal' data. Since most `BraTS` volumes contain tumors (anomalies), training data will be carefully selected. We will explore two main strategies: Training exclusively on volumes identified as "healthy controls" if available and sufficient in the dataset subset used.

Training on volumes containing tumors but extracting only non-tumorous regions (using ground truth masks) for the training set. This approach requires careful handling to ensure spatial context is maintained (e.g., training on small 3D patches or 2D slices confirmed to be tumor-free). The latter strategy is more likely given the nature of `BraTS`, but requires mask usage only for training data preparation and evaluation, never for training the detection model itself. The test set will comprise volumes containing tumors, evaluated on their ability to identify the anomalous regions.

3.2 `GANomaly` Architecture Adaptation and Training

Architecture: We will implement an adapted version of the GANomaly architecture, as described in [10] and reviewed in [11].
This architecture consists of:

- **Generator (G)** with an autoencoder structure (Encoder GE, Decoder GD, and a second Encoder E). GE maps input image x to latent space z. GD reconstructs x from z to x'. E maps the reconstructed image x' back to the latent space z'.
- **Discriminator (D)** trained to distinguish between real images x and generated images $x' = G(z)$, and also potentially between pairs (x, z) and (x', z') depending on the exact variant. Following the standard GANomaly, D will output a prediction indicating whether the input is real or fake, and extract feature representations.
- **Adaptation for BraTS:** The architecture will be adapted for processing 2D slices extracted from the 3D BraTS volumes or using 3D convolutional layers for volumetric patch processing, depending on the chosen data preparation strategy. Filter sizes, strides, and network depth will be adjusted accordingly.
- **Unsupervised Training:** The model will be trained exclusively on the prepared 'normal' data. The training objective aims to minimize a combined loss function for the Generator while the Discriminator is trained adversarially.

The Generator's loss function will combine:

- **Adversarial Loss ($\mathbf{L}_{adv}$):** Encourages the generator to produce realistic images, often using feature matching from the discriminator [8].
- **Contextual Loss ($\mathbf{L}_{con}$):** Measures the pixel-wise or spatial similarity between the input image x and its reconstruction x' (e.g., ℓ_1 loss).
- **Encoder Loss ($\mathbf{L}_{enc}$):** Encourages the autoencoder part of the generator to map the reconstructed image x' back to a latent representation z' similar to the original latent mapping z from the input x (e.g., ℓ_2 loss on z and z').

Training will involve alternating updates for the Discriminator and the Generator, using optimizers like Adam, over a defined number of epochs.

3.3 Anomaly Detection and Scoring (Novelty Exploration)

Once trained on 'normal' data, the `GANomaly` model is used for inference on unseen test volumes containing anomalies. An anomalous region is expected to deviate significantly from the learned 'normal' data distribution, resulting in a high anomaly score.
The standard `GANomaly` approach uses the latent space difference (Encoder Loss, $\mathcal{L}_{\text{enc}}$) between the latent representation of the input image (z) and the latent representation of its reconstruction (z') as the anomaly score [10]. The intuition is that anomalous inputs will be poorly reconstructed, leading to a large difference in latent space representation.
Novelty: Building on the different scoring mechanisms discussed in the survey [11] (which covers reconstruction error in `AnoGAN` and latent space differences in `GANomaly`), this project will explore the comparative effectiveness of different anomaly scoring methods derived from the trained `GANomaly` model on the BraTS data. Specifically, we will compare: Using the Encoder Loss ($\mathcal{L}_{\text{enc}}$) as the primary anomaly score, as in the original `GANomaly`.
Using the Contextual Loss ($\mathcal{L}_{\text{con}}$), i.e., the reconstruction error between the input x and reconstructed x', as the anomaly score. This aligns more closely with the `AnoGAN` scoring principle but applied within the `GANomaly` framework.
Evaluating a weighted combination of $\mathcal{L}_{\text{enc}}$ and $\mathcal{L}_{\text{con}}$ as the anomaly score (e.g., "Score = $\alpha \cdot \mathcal{L}_{\text{enc}} + \beta \cdot \mathcal{L}_{\text{con}}$").
For each test image (slice/patch), these different scores will be computed pixel-wise or patch-wise, generating an anomaly map.

3.4 Evaluation Methodology

The performance of the anomaly detection system using each of the proposed scoring methods (L_{enc}, L_{con}, Weighted Combination) will be evaluated against the ground truth tumor masks provided in the `BraTS` dataset.
Evaluation metrics will include: Area Under the Receiver Operating Characteristic Curve (`AUC-ROC`) or Area Under the Precision-Recall Curve (`AUPRC`): These will be computed by treating each pixel/`voxel`/patch in the test set as a sample, comparing its anomaly score (after appropriate scaling) against whether it falls within a ground truth tumor region or not. AUPRC is particularly suitable for imbalanced datasets where the anomalous class is small, as is typical for tumors within a large brain volume. The survey paper [11] used `AUPRC`, making it a relevant choice for comparison with existing literature on other datasets.
Spatial Overlap Metrics (e.g., Dice Similarity Coefficient): An anomaly score threshold will be determined (e.g., empirically on a separate validation subset or by selecting a threshold based on

ROC/PR curves). This threshold will be used to generate a binary anomaly detection mask. The Dice coefficient will then be calculated between this detected mask and the ground truth tumor mask for each test volume/slice. This provides a spatial measure of detection accuracy.
Quantitative comparison using these metrics will determine which anomaly scoring strategy performs best for detecting brain tumors within the `GANomaly` framework on the `BraTS` dataset.

3.5 Software Design, Development & Tools

The project will primarily involve software development for implementing the `GANomaly` architecture, data preprocessing pipeline, anomaly scoring logic, and evaluation procedures.
Tools: The implementation will leverage Python with a deep learning framework (e.g., PyTorch or TensorFlow). Libraries for medical image processing (e.g., `SimpleITK`, `NiBabel`), numerical operations (`NumPy`, `SciPy`), and data manipulation (Pandas) will also be used. Computationally intensive training will require access to GPU resources.
Development: An existing open-source implementation of `GANomaly` (if available and compatible, potentially adapting the one mentioned in [11] if in Python/TensorFlow) will serve as a starting point. The core development will focus on:

- Adapting the network architecture for 2D slices or 3D data.
- Implementing the robust`BraTS` preprocessing pipeline.
- Implementing the different anomaly scoring calculation methods (`Lenc`,`Lcon`, combined).

Developing the evaluation scripts using the defined metrics and comparing the different scores. Evaluation of Software: The correctness of the implementation will be verified through unit testing of individual components (e.g., preprocessing steps, loss calculations, scoring functions). Model training progress will be monitored using standard metrics (loss curves, discriminator/-generator balance). Final evaluation is based on the anomaly detection performance metrics as described in Section 3.4.

4 Work Plan

The work plan for this project is presented in Figure3, which outlines the major tasks week-by-week over the planned project timeline. Each task has been defined with a realistic time estimate, dependencies, and clearly marked milestones. Tasks are grouped into five main stages: *Project Setup & Review, Data Handling, Model Implementation, Evaluation, and Reporting & Finalization*.

To ensure a structured and goal-driven progression, the work plan is aligned with the project objectives and integrates overlapping work packages when feasible. A buffer for potential delays is embedded within the evaluation and reporting phases to allow flexibility without compromising deadlines.

5 Risks

This research project, while promising, is subject to several potential risks that could impact its progress and outcomes. To proactively manage these challenges, a risk register has been developed, outlining each identified risk, its likelihood, its potential impact, proactive mitigation

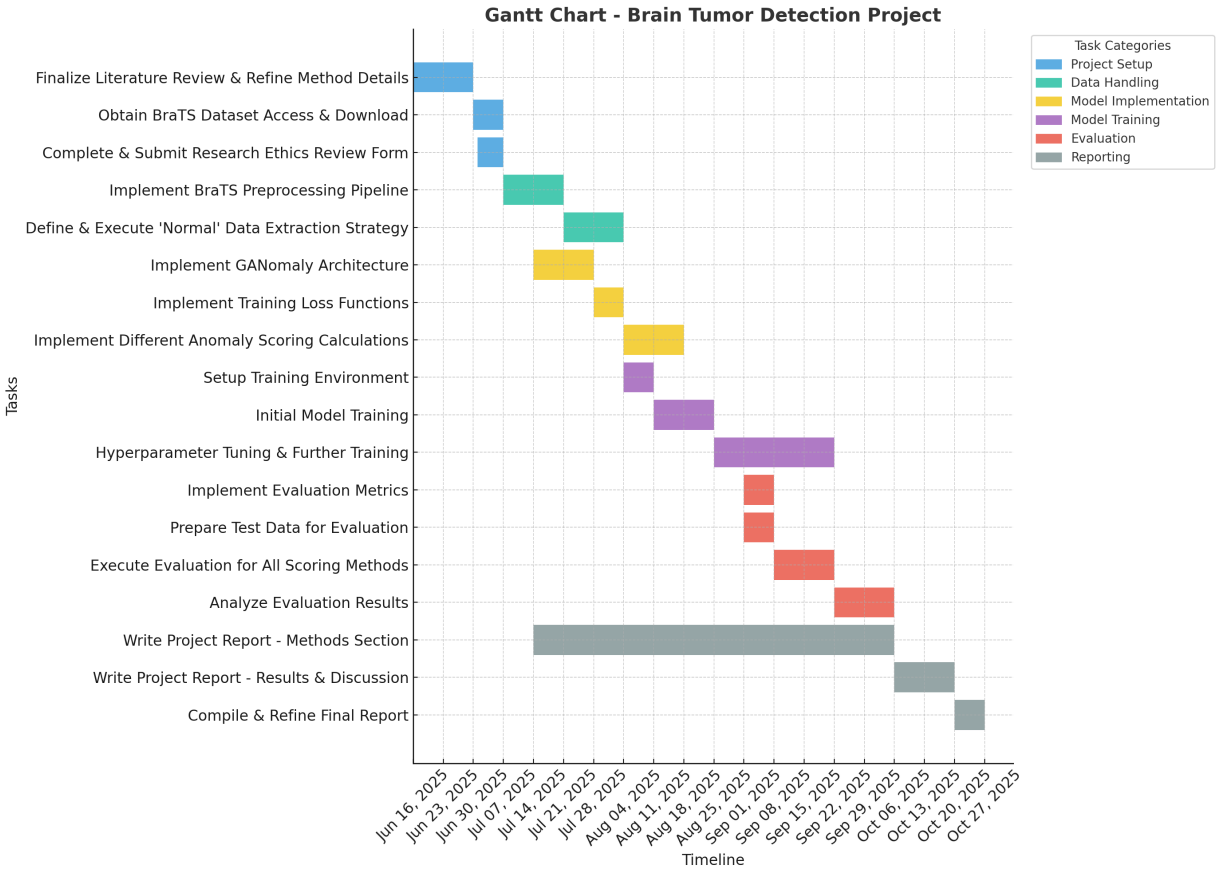

Figure 3: Gantt chart showing the timeline of major project tasks and milestones.

strategies to reduce the probability of the risk occurring, and recovery plans should the risk materialize, with the goal of minimising its impact. This proactive approach ensures that potential disruptions are anticipated and addressed. The table 1 presents the risk register.

6 Presentation

The project proposal is presented as a coherent and professional document. Care has been taken to ensure a clear narrative flow between sections, guiding the reader logically from the project's introduction and context through the detailed methodology, work plan, and risk assessment. The language is clear, concise, and appropriate for an academic proposal. Formatting is consistent throughout, including headings, text styles, and figure/table captions. All figures, particularly the graphical work plan, are designed to be legible and informative. Academic conduct is strictly maintained through accurate in-text citation and a comprehensive reference list, formatted according to the Harvard style. The document has been thoroughly proofread to minimise grammatical errors and typos, ensuring a professional and polished submission.

7 Ethical, Legal & Professional Issues

Critical ethical, legal, and professional considerations are paramount for this research project involving sensitive medical imaging data. As detailed comprehensively within Section 3 (Approaches: Methods & Tools), the project adheres to strict protocols for data privacy and security,

Table 1: Risk register detailing identified risks, their likelihood, impact, and mitigation strategies.

Risk	**Likelihood** (Low/Medium/High)	**Impact** (Low/Medium/High)	**Mitigation Strategy**
Data Availability and Quality	Medium	High	Proactive: Utilize the BraTS dataset for access and focus on the BraTS dataset. Explore the dataset to determine if the dataset is free of outliers.
Computational Resource Constraints	Medium	Medium	Proactive: Start research early, and carefully select model architecture. Plan training schedules, and discuss needs with the supervisor.
Model Convergence Issues	Medium	Medium	Proactive: Carefully select the loss functions, and apply the appropriate optimizer. Optimize the hyperparameters with the validation set.
Data Imbalance and Limited Data for Training	High	Medium	Proactive: Data augmentation techniques, such as random rotations, flipping, and intensity adjustments. Address potential data imbalance.
Difficulty in Achieving State-of-the-Art Performance	Medium	Medium	Proactive: Implement comprehensive evaluation metrics. Thoroughly review related papers.

respects consent and data usage restrictions of the `BraTS` dataset, acknowledges and discusses potential biases inherent in the data, and commits to responsible development and communication of research findings. The mandatory requirement for addressing these issues is recognised, and the detailed discussion in Section 3 serves as the foundation for the ethical framework guiding this project. These considerations have also informed the completion of the mandatory Research Ethics Review Form.

References

[1] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Proceedings of the 27th International Conference on Neural Information Processing Systems (NeurIPS)*, pages 2672–2680, Montreal, QC, Canada, 2014. Curran Associates, Inc.

[2] The BraTS 2021 Challenge Organizers. Rsna-asnr-miccai brain tumor segmentation (brats) challenge 2021. `https://www.kaggle.com/datasets/dschettler8845/brats-2021-task1`, 2021. BraTS2021: The Brain Tumor Segmentation (BraTS) challenge's 10th anniversary, organized by RSNA, ASNR, and MICCAI. Focused on mpMRI-based glioblastoma segmentation and MGMT promoter methylation status classification.

[3] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015. URL `https://doi.org/10.48550/arXiv.1511.06434`.

[4] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets, 2014. URL `https://arxiv.org/abs/1411.1784`.

[5] Saddam Hussain, Syed Muhammad Anwar, and Muhammad Majid. Segmentation of glioma tumors in brain using deep convolutional neural network. *Neurocomputing*, 282:248–261, 2018. ISSN 0925-2312. doi: https://doi.org/10.1016/j.neucom.2017.12.032. URL `https://www.sciencedirect.com/science/article/pii/S0925231217318763`.

[6] Likun Cai, Yanjie Chen, Ning Cai, Wei Cheng, and Hao Wang. Utilizing amari-alpha divergence to stabilize the training of generative adversarial networks. *Entropy*, 22(4), 2020. ISSN 1099-4300. doi: 10.3390/e22040410. URL `https://www.mdpi.com/1099-4300/22/4/410`.

[7] Abhishta Bhandari, Jarrad Koppen, and Marc Agzarian. Convolutional neural networks for brain tumour segmentation. *Insights into Imaging*, 11(1):77, 2020. doi: 10.1186/s13244-020-00892-7. URL `https://doi.org/10.1186/s13244-020-00892-7`.

[8] Thomas Schlegl, Philipp Seeböck, Sebastian M. Waldstein, Ursula Schmidt-Erfurth, and Georg Langs. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In *International Conference on Information Processing in Medical Imaging (IPMI)*, volume 10265 of *Lecture Notes in Computer Science*, pages 146–157. Springer, 2017. doi: 10.1007/978-3-319-59050-9_12. URL `https://doi.org/10.1007/978-3-319-59050-9_12`.

[9] Shobhita Sundaram and Neha Hulkund. Gan-based data augmentation for chest x-ray classification. *arXiv preprint arXiv:2007.08029*, 2020. URL `https://arxiv.org/abs/2007.08029`.

[10] Samet Akcay, Amirghassem Abarghouei, and Toby P. Breckon. GANomaly: Semi-Supervised Anomaly Detection via Adversarial Training. *arXiv preprint arXiv:1805.06725*, 2018. URL `http://arxiv.org/abs/1805.06725`.

[11] Federico Di Mattia, Paolo Galeone, Michele De Simoni, and Emanuele Ghelfi. A Survey on GANs for Anomaly Detection. *arXiv preprint arXiv:1906.11632*, 2019. URL `https://arxiv.org/abs/1906.11632`.

Research Ethics Review Form: BSc, MSc and MA Projects

Computer Science Research Ethics Committee (CSREC)

http://www.city.ac.uk/department-computer-science/research-ethics

Undergraduate and postgraduate students undertaking their final project in the Department of Computer Science are required to consider the ethics of their project work and to ensure that it complies with research ethics guidelines. In some cases, a project will need approval from an ethics committee before it can proceed. Usually, but not always, this will be because the student is involving other people ("participants") in the project.

In order to ensure that appropriate consideration is given to ethical issues, all students must complete this form and attach it to their project proposal document. There are two parts:

PART A: Ethics Checklist. All students must complete this part.
The checklist identifies whether the project requires ethical approval and, if so, where to apply for approval.

PART B: Ethics Proportionate Review Form. Students who have answered "no" to all questions in A1, A2 and A3 and "yes" to question 4 in A4 in the ethics checklist must complete this part. The project supervisor has delegated authority to provide approval in such cases that are considered to involve MINIMAL risk.
The approval may be ***provisional*** – *identifying the planned research as* likely to involve MINIMAL RISK.
In such cases you must additionally seek ***full approval*** from the supervisor as the project progresses and details are established. ***Full approval*** must be acquired in writing, before beginning the planned research.

A.1 If you answer YES to any of the questions in this block, you must apply to an appropriate external ethics committee for approval and log this approval as an External Application through Research Ethics Online - https://ethics.city.ac.uk/		*Delete as appropriate*
1.1	Does your research require approval from the National Research Ethics Service (NRES)? *e.g. because you are recruiting current NHS patients or staff?* *If you are unsure try - https://www.hra.nhs.uk/approvals-amendments/what-approvals-do-i-need/*	**NO**
1.2	Will you recruit participants who fall under the auspices of the Mental Capacity Act? *Such research needs to be approved by an external ethics committee such as NRES or the Social Care Research Ethics Committee - http://www.scie.org.uk/research/ethics-committee/*	**NO**
1.3	Will you recruit any participants who are currently under the auspices of the Criminal Justice System, for example, but not limited to, people on remand, prisoners and those on probation? *Such research needs to be authorised by the ethics approval system of the National Offender Management Service.*	**NO**
A.2 If you answer YES to any of the questions in this block, then unless you are applying to an external ethics committee, you must apply for approval from the Senate Research Ethics Committee (SREC) through Research Ethics Online - https://ethics.city.ac.uk/		*Delete as appropriate*
2.1	Does your research involve participants who are unable to give informed consent? *For example, but not limited to, people who may have a degree of learning disability or mental health problem, that means they are unable to make an informed decision on their own behalf.*	**NO**
2.2	Is there a risk that your research might lead to disclosures from participants concerning their involvement in illegal activities?	**NO**
2.3	Is there a risk that obscene and or illegal material may need to be accessed for your research study (including online content and other material)?	**NO**

2.4	Does your project involve participants disclosing information about special category or sensitive subjects? *For example, but not limited to: racial or ethnic origin; political opinions; religious beliefs; trade union membership; physical or mental health; sexual life; criminal offences and proceedings*	**NO**
2.5	Does your research involve you travelling to another country outside of the UK, where the Foreign & Commonwealth Office has issued a travel warning that affects the area in which you will study? *Please check the latest guidance from the FCO - http://www.fco.gov.uk/en/*	**NO**
2.6	Does your research involve invasive or intrusive procedures? *These may include, but are not limited to, electrical stimulation, heat, cold or bruising.*	**NO**
2.7	Does your research involve animals?	**NO**
2.8	Does your research involve the administration of drugs, placebos or other substances to study participants?	**NO**
A.3 If you answer YES to any of the questions in this block, then unless you are applying to an external ethics committee or the SREC, you must apply for approval from the Computer Science Research Ethics Committee (CSREC) through Research Ethics Online - https://ethics.city.ac.uk/ **Depending on the level of risk associated with your application, it may be referred to the Senate Research Ethics Committee.**		*Delete as appropriate*
3.1	Does your research involve participants who are under the age of 18?	**NO**
3.2	Does your research involve adults who are vulnerable because of their social, psychological or medical circumstances (vulnerable adults)? *This includes adults with cognitive and / or learning disabilities, adults with physical disabilities and older people.*	**NO**
3.3	Are participants recruited because they are staff or students of City, University of London? *For example, students studying on a particular course or module.* *If yes, then approval is also required from the Head of Department or Programme Director.*	**NO**
3.4	Does your research involve intentional deception of participants?	**NO**
3.5	Does your research involve participants taking part without their informed consent?	**NO**
3.5	Is the risk posed to participants greater than that in normal working life?	**NO**
3.7	Is the risk posed to you, the researcher(s), greater than that in normal working life?	**NO**
A.4 If you answer YES to the following question and your answers to all other questions in sections A1, A2 and A3 are NO, then your project is deemed to be of MINIMAL RISK. **If this is the case, then you can apply for approval through your supervisor under PROPORTIONATE REVIEW. You do so by completing PART B of this form.** **If you have answered NO to all questions on this form, then your project does not require ethical approval. You should submit and retain this form as evidence of this.**		*Delete as appropriate*
4	Does your project involve human participants or their identifiable personal data? *For example, as interviewees, respondents to a survey or participants in testing.*	**NO**