Development of an Automated System for Diagnosis of Ductal Carcinoma using Deep Learning Techniques

Thesis submitted by

Soumyajyoti Dey

Doctor of Philosophy(Engineering)

Department of Computer Science and Engineering
Faculty Council of Engineering & Technology
Jadavpur University
Kolkata, India

2025

JADAVPUR UNIVERSITY
KOLKATA, INDIA

Index No: 237/20/E

1. **Title of the Thesis:**
   Development of an Automated System for Diagnosis of Ductal Carcinoma using Deep Learning Techniques

2. **Name, Designation & Institution of the Supervisor**

   Dr. Nibaran Das,
   Professor, Department of Computer Science & Engineering
   Jadavpur University, Kolkata- 700032, India

3. **List of Publications:**

   (a) Journal Publications

      i. **S. Dey**, S. Mitra, S. Chakraborty, D. Mondal, M. Nasipuri, and N. Das, "GC-EnC: A Copula based ensemble of CNNs for malignancy identification in breast histopathology and cytology images", Computers in Biology and Medicine, vol. 152, p. 106329, 2023. (impact factor: 7.0)

      ii. **S. Dey**, S. Mitra, S. Chakraborty, M. Nasipuri, & N. Das, "BrCyto-GAN:Fuzzy Entropy based Selective Augmentation from the Hierarchical GAN generated Super Resolution Breast Cytology Images". (*Communicated*)

   (b) Conference Publications

      i. **S. Dey**, S. Chakraborty, U. G. Roy, and N. Das, "Could We Generate Cytology Images from Histopathology Images? An Empirical Study", in Proceedings of Third International Conference on Advanced Computing and Applications, 2024, pp. 625–635.

      ii. **S. Dey**, S. Chakraborty, U. G. Roy, and N. Das, "Fuzzy Rank-Based Late Fusion Technique for Cytology Image Segmentation", in Proceedings of International Conference on Data, Electronics and Computing, 2024, pp. 11–23.

iii. **S. Dey**, S. Nasipuri, O. Ghosh, S. Chakraborty, D. Mondal, and N. Das, "Malignancy Identification from Cytology Images Using Deep Optimal Features", in Proceedings of International Conference on Data, Electronics and Computing, 2023, pp. 279–287.

iv. **S. Dey**, S. Das, S. Ghosh, S. Mitra, S. Chakrabarty, and N. Das, "SynCGAN: Using Learnable Class Specific Priors to Generate Synthetic Data for Improving Classifier Performance on Cytological Images", in Computer Vision, Pattern Recognition, Image Processing, and Graphics, 2020, pp. 32–42.

4. **List of Patents:** None

5. **List of Presentations in National/International Conferences**

i. **S. Dey**, S. Chakraborty, U. G. Roy, and N. Das, "Could We Generate Cytology Images from Histopathology Images? An Empirical Study", in Proceedings of Third International Conference on Advanced Computing and Applications, 2024, pp. 625–635.

ii. **S. Dey**, S. Chakraborty, U. G. Roy, and N. Das, "Fuzzy Rank-Based Late Fusion Technique for Cytology Image Segmentation", in Proceedings of International Conference on Data, Electronics and Computing, 2024, pp. 11–23.

iii. **S. Dey**, S. Mitra, S. Chakraborty, D. Mondal, M. Nasipuri, and N. Das, "GC-EnC: A Copula based ensemble of CNNs for malignancy identification in breast histopathology and cytology images", Computers in Biology and Medicine, vol. 152, p. 106329, presented at 18^{th} Academic Research and Careers for Students Symposium(ACM INDIA ARCS-2024), NISER Bhubaneswar.

Statement of Originality

I, **Soumyajyoti Dey** registered on **17/02/2020** do hereby declare that this thesis entitled **"Development of an Automated System for Diagnosis of Ductal Carcinoma using Deep Learning Techniques"** contains literature survey and original research work done by the undersigned candidate as part of Doctoral studies.

All information in this thesis have been obtained and presented in accordance with existing academic rules and ethical conduct. I declare that, as required by these rules and conduct, I have fully cited and referred all materials and results that are not original to this work.

I also declare that I have checked this thesis as per the "Policy on Anti Plagiarism, Jadavpur University, 2019", and the level of similarity as checked by iThenticate software is .3.%.

Signature of the Candidate : Soumyajyoti Dey

Date : 13/5/2025

Certified by Supervisor:
(Signature with date & seal)

Nibaran Das
13/5/25
Professor
Computer Sc. & Engg. Department
Jadavpur University
Kolkata - 700032

1.

CERTIFICATE FROM THE SUPERVISOR

This is to certify that the thesis entitled **"Development of an Automated System for Diagnosis of Ductal Carcinoma using Deep Learning Techniques"** submitted by **Shri Soumyajyoti Dey**, who got his name registered on **17/02/2020** for the award of Ph. D. (Engg.) degree of Jadavpur University is absolutely based upon his/her own work under the supervision of **Prof. Nibaran Das** and that neither his thesis nor any part of the thesis has been submitted for any degree/diploma or any other academic award anywhere before.

Nibaran Das. 13/5/25

Dr. Nibaran Das
Professor
Department of Computer Science and Engineering
Jadavpur University

Professor
Computer Sc. & Engg. Department
Jadavpur University
Kolkata - 700032

ACKNOWLEDGMENTS

Pursuing this Ph.D. has profoundly transformed my perspective on life. Over the past few years, I have been fortunate to receive unwavering support and guidance from many remarkable individuals.

First and foremost, I would like to express my heartfelt gratitude to my supervisor, Dr. Nibaran Das, Professor, Department of Computer Science and Engineering, Jadavpur University, for his invaluable guidance, insightful suggestions, continuous support, patience, motivation, and unwavering faith in me. Thank you for leading me through this incredible journey.

My sincere thanks to Prof. Nirmalya Chowdhury, Head of the Department of CSE, Jadavpur University, for providing all necessary departmental facilities that enabled me to carry out my research smoothly.

I am also grateful to the Centre for Microprocessor Applications for Training, Education and Research (CMATER), Department of CSE, Jadavpur University, for granting access to essential facilities and infrastructure that significantly supported my research endeavors.

I would like to extend my sincere appreciation to Dr. Mita Nasipuri (Former Professor, Dept. of CSE, Jadavpur University) and Dr. Shyamali Mitra (Assistant Professor, Dept. of IEE, Jadavpur University) for offering valuable insights that helped broaden my research perspective across related domains.

This Ph.D. journey would not have been possible without financial support. I express my gratitude to the Science and Engineering Research Board (SERB), Government of India, for funding my research through the project (Ref. No: EEQ/2018/000963), SVMCM, Govt. of WB, and to the Jadavpur University Alumni Association (Bombay Branch) Trust for awarding me the LATE A.K. Basu Research Scholarship. I am also thankful to ACM India for granting me the Anveshan Setu Fellowship–2023, which allowed me to work at the Visual Computing Lab, IISc Bangalore, under the mentorship of Dr. Anirban Chakraborty (Associate Professor, CDS Department, IISc Bangalore).

I sincerely thank Theism Medical Diagnostics Centre, Kolkata, West Bengal, for providing the cytological sample slides crucial to my Ph.D. work. I am especially grateful to Dr. Sukanta Chakraborty, Dr. Debashri Mondal, and Dr. Utso Guha Roy for their continued guidance and support in data collection and analysis.

To all my co-authors—thank you for being a part of this amazing journey. I also extend my thanks to my fellow researchers, friends, and seniors from the CMATER Laboratory, Jadavpur University—Somenath Kuiry, Rahul Laxmanrao Meshram, Dibyasree Guha, Raju Naskar, Snehasis Sahu, Pabitra Kumar Mondal, Swarnendu Ghosh, Kalpita Dutta, and Bidhan Barai—for their constant encouragement and assistance throughout this journey.

Last but not least, my deepest gratitude goes to my parents, whose unwavering love, support, and belief in me have been my greatest strength.

Date: 13/5/2025

Soumyajyoti Dey

Soumyajyoti Dey

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Cancer, as of now, has become a soaring concern among individuals. It is also one of the most dreadful diseases in today's life. According to WHO (World Health Organization) 2022 report, 20 million new cancer cases [26] were found globally, and 9.7 million patients have lost their lives to cancer. Traditionally living organisms are composed of a large number of cells. Normal cells have a specific lifetime in which they undergo division to produce new cells, replacing old ones. Before the old cell is worn out, it passes its genetic information to new ones. But cancer cells divide in a proliferative manner disobeying coordination among cells. Though cancer-causing genes are genetically inherited, various external factors like indulging in smoking and drinking, exposure to heavy metals, radiation, usage of plastics, and so on, are also responsible for the growth of cancer cells [61]. It is one of the biggest challenging areas in medical science because its cure depends on the correct diagnosis at the right stage. Also, the earlier and more accurate diagnoses can reduce the mortality rate. Traditional manual diagnosis is time-consuming and may lead to delays in treatment. Therefore, automation plays a crucial role in improving efficiency and accuracy. Among various types of cancer, the most common ones among Indian women are breast, cervical, respiratory tract, ovarian, and colorectal cancers. This study focuses on the automated diagnosis of breast cancer using advanced computer vision techniques.

1.1 Ductal Carcinoma of Breast

Breast cancer, though mostly prevalent among women all over the world, occurrences found in men are rare (cases found are mostly invasive in nature). In women, the incidence and deaths due to breast cancer is the highest among all types of cancer [102], and thus posing a global challenge irrespective of all levels of modernization. Breast cancers originate either in the lobules or in the ducts that connect lobules to the nipple. Breast cancer generally does not show symptoms (before stage III or IIIc) when the tumor is small enough to be felt and can be easily cured. It starts as a painless lump and slowly progresses with other symptoms like pain, breast heaviness, swelling, and redness of the skin. Nipple abnormalities such as spontaneous watery or bloody discharge or retraction are also common. There are several reasons for suffering breast cancers like age, family history, genetic mutation, obesity and alcohol consumption, etc.

Breast cancer is generally classified into two main types: (i) Carcinoma In Situ, and (ii) Invasive [10].

(i) Carcinoma In Situ: This refers to the presence of abnormal cells that remain confined to their place of origin and have not invaded surrounding cells. There are two primary types: ductal carcinoma in situ (DCIS) and lobular carcinoma in situ (LCIS).

- Ductal carcinoma in situ [20]: In this condition, abnormal cells replace the normal epithelial cells lining the breast ducts. However, it does not necessarily progress to an invasive stage.
- In this condition, abnormal cell growth is confined to certain lobules of the breast and may progress to invasive cancer.

(ii) Invasive: It has two subtypes: (a) Regional stage: Abnormal cells spread to adjacent tissues and nearby lymph nodes, typically corresponding to stage II or III breast cancer. (b) Distant stage: A condition in which abnormal cells have metastasized to different organs or lymph nodes above collarbone (stage IIIc and stage IV cancers).

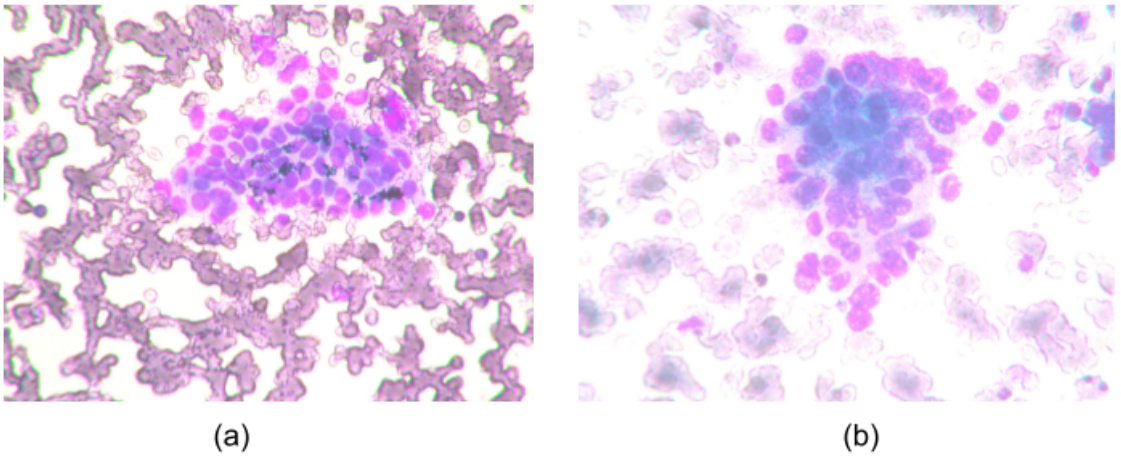


Figure 1.1: Sample images of Breast cytology (40× magnification): (a) Benign case, (b) Malignant case

For the diagnosis of ductal carcinoma, the practitioners are using different diagnostic techniques like mammography, ultrasound, MRI, biopsy, immunohistochemistry tests etc. Among all the most common is the biopsy technique. There are different kinds of biopsy techniques, like Minimally Invasive Biopsy Techniques (Fine Needle Aspiration Biopsy(FNAC), Core Needle Biopsy) and Surgical Biopsy Techniques. In this study, we have mainly focused on cytology samples obtained through FNAC test. Histopathology samples, obtained by tissue biopsy, are also explored for validating the automated model performances.

1.1.1 Cytology Image

The body of an organism is composed of trillions of cells. Each cell possesses a cytoplasm and a nucleus. Cytoplasm, or the cell body, acts as an envelope to the nucleus containing chromosomes, the genetic material that undergoes mutation under certain changes. The changes are reflected in the morphology of the nucleus and cytoplasm. Thus, a micro examination of cells can unfold relevant information on any morphological alteration in response to a particular disease. These cells, which are important predictors of pre-malignant and malignant lesions can be sampled and examined under a microscope to diagnose different medical conditions except a few. FNAC is a procedure to primarily assess the cause of swelling by examining cells in it. The doctor who performs the process usually fixes the swelling with one hand and introduces a 24-gauge needle with a 10-ml syringe attached to it. By creating a negative suction the needle is moved within

the swelling. The cellular aspirate comes within the hub of the needle. It is then expressed on glass slides and smears are made with another glass slide called spreader. The slide containing material is then processed by two methods either Air drying and staining with May-Grunwald-Giemsa (MGG) stain or alcohol fixation using Papanicolau or Pap stain. This process needs expertise as exact swelling is to be targeted without damaging the surrounding vital structure. Also, the cellular material should come out with less admixture of blood. FNAC is effective for the primary evaluation of any swelling as well as a safe procedure with very little pain. So, the branch of medical science that deals with the examination of cells to detect the presence of any abnormal cellularity is known as cytology.

The cellular characteristics of benign and malignant samples are depicted as follows in details:

Cellular Characteristics of Cytology Specimen

- **Benign Sample:** Benign cytology samples exhibit some morphological characteristics, which indicates the non-cancerous condition.

  1. **Nuclear Features:** Small, inconspicuous nucleoli [49] are present. There is no evidence of mitotic figures or hyperchromasia.
  2. **Cytoplasmic Features:** The cytoplasms are in moderate amount and well-defined. The cytoplasmic borders are distinct, and the staining is uniform [24].
  3. **Background Characteristics:** Abundant stromal fragments with a fibromyxoid appearance, often associated with fibroadenoma [84]. The stroma appears hypocellular or contains scattered stromal elements. No necrotic debris or tumor diathesis is seen. No significant nuclear atypia, pleomorphism, or mitotic activity is observed.
  4. **Epithelial Cells:** Appear as cohesive clusters, monolayered sheets, or branching antler-like (stromal) structures. Cells are uniform with a bland cytological appearance, well-preserved cytoplasm and a low nuclear to cytoplasmic (N:C) ratio [5]. Nuclei are round to oval, with smooth contours and fine, evenly distributed chromatin.

  5. **Stromal Cells:** It follows the biphasic pattern. Both epithelial and stromal components are present in fibroadenomas[104]. Spindle-shaped stromal cells may be seen, often arranged loosely.

Malignant Sample: The cytology sample has been classified as malignant with respect to some morphological characteristics. The following characteristics are present in the ductal carcinoma sample:

- **Cell Size and Shape:** Marked variation in cell size and shape [83]. Some cells may appear unusually large or irregular.
- **Nuclear Features:** Nuclei show significant variation in size and shape. Nuclei are darkly stained, reflecting increased DNA content. Chromatin is clumped, irregular, and unevenly distributed. One or more prominent, irregular nucleoli and increased nuclear-to-cytoplasmic ratio due to nuclear enlargement are commonly seen.
- **Cytoplasmic Features:** Cytoplasm may appear scant and poorly defined. Cytoplasm often stains darkly due to increased ribosomal activity.
- **Background Features:** The background is dirty or necrotic with cellular debris [3]. Background may show an associated inflammatory infiltrate, including lymphocytes or macrophages.
- **Arrangement of Cells:** Cells are typically seen in loosely cohesive clusters, sheets, or as singly dispersed cells. Rarely, cells may form duct-like structures, particularly in Ductal carcinoma in situ(DCIS).

Some examples of benign and malignant sample are shown in Figure 1.1.

1.2 Deep Learning based Automation in Ductal Carcinoma diagnosis

In the recent advancement of computer vision techniques, automated diagnosis of breast cancer from cytology or histopathology images is giving significant and accurate performance. Earlier, various machine learning based computer vision techniques [51, 120, 73] were applied for breast cancer diagnosis, but they required domain expertise to extract handcrafted features. However, with the introduction of deep learning, particularly Convolutional Neural Networks (CNNs), features

are automatically extracted through multiple hidden layers, enabling more efficient automated prediction. Several deep learning-based approaches have already been reported for malignancy identification in breast cytology images. Saikia et al. [89] developed a CNN-based deep learning classification framework where images were augmented using techniques like cropping, shearing, rotation, mirroring, skewing, inverting, zooming. In subsequent phases, channel identification, histogram equalization and Otsu's thresholding were used to segment candidate nuclei. Maximum accuracy of 96.25% was achieved using GoogLeNet architecture. Kowal et al. [57] proposed a CNN and seeded watershed-based breast cytology segmentation model. CNN-based semantic segmentation model was first applied to differentiate between nuclei and background. The generated semantic mask was transformed into a nuclei mask to extract touching and overlapping nuclei. The clustered nuclei were detected by area and roundness. Nuclei seeds were identified using conditional erosion process. The overlapping nuclei were separated by seeded watershed algorithm. With this approach, 83.4% of benign nuclei were classified using Hausdorff distance. Park et al. [79] proposed a deep learning based automated system for metastatic breast cancer diagnosis. In this work, 569 cytological slides are used for malignant pleural effusion. The samples are classified by deep CNN model Inception-ResNet-V2 and achieved accuracy, specificity 87.9% and 95.7% respectively. The concept of transfer learning is implemented by Khan et al. [52] for the classification of breast cytology images. Here, firstly the features are extracted from pre-trained GoogLeNet, VGGNet, and ResNet architectures, and it are fed into fully connected layer using average pooling. The proposed framework has achieved average accuracy 97.52%.

1.3 Motivation

Digital pathology image analysis, particularly cytology imaging, plays a crucial role in the early diagnosis of breast cancer, which is essential for effective treatment and improved patient survival. Traditionally, cytology images have been analyzed through visual interpretation for accurate diagnosis, with cellular structures examined under a microscope by skilled practitioners. However, manual annotation is time-consuming and requires significant expertise. Additionally, pathologists often face an overwhelming workload as breast cancer cases continue to rise at an

alarming rate. To address these challenges, Computer-Aided Diagnosis (CAD) systems have been developed to assist in breast cancer diagnosis by reducing report generation time, alleviating the workload of cytologists, and minimizing human errors. In recent years, advancements in computer vision techniques have significantly impacted digital pathology image analysis, particularly for cancer diagnosis. Deep Convolutional Neural Network (DCNN) models have demonstrated remarkable performance in cytology image segmentation, localization, feature extraction, classification. Deep learning-based CAD systems can provide real-time decision support for pathologists and deliver highly accurate results.

However, despite these advancements, existing DCNN models are not fully fine-tuned for cytology image analysis due to several challenges. A major issue is the limited availability of adequately annotated breast cytology samples, as annotation is expensive, requires expert knowledge, and must adhere to ethical guidelines. Also, the presence of publicly available breast cytology image datasets is quite limited, although they are essential for training and validating deep learning models. The performance of deep leaning models, like CNNs, highly depends on the data size. If the data size is small, CNNs overfit the data resulting in reduction of their generalization ability. A model's generalization ability ensures the good performance of the model on both familiar (the training data) and unknown situations (the test data). Additionally, cytology image datasets are often imbalanced, with fewer malignant samples compared to normal ones, leading to biased models. To solve this kind of problem, we can focus on some synthetic data generation techniques. But the generation of synthetic sample from noise is quite challenging due to complex cellular structures. The generative models can also suffer from mode collapse, which can produce limited variations of samples. But for developing a robust classification system, we need a wide variation of training samples for improving the learning performance of CNN model.

In the cytology samples, there are some ovarlapping regions among nuclei or cytoplasm, and sometimes we can't find the proper cellular structure for uneven or excessive spreading of staining meterial during slide preparation. So it can significantly affect the segmentation task, and as a result we can find false positive or false negative cases at the time of classification. In the cytology samples, there are several cellular objects in different shapes and orientation, but sometimes different objects form shape of another object. For automated localization of cellular objects, it can detect some wrong cellular objects. Sometimes, localization and

segmentation performance is limited due to the lack of pixel-level annotations, as most standard datasets provide labels only at the patch or region-of-interest (ROI) level.

Ensemble techniques are well-known for improving classification performance by combining multiple models; however, selecting a suitable ensemble method for a specific domain remains a significant challenge. The features extracted from CNN models may not always yield good classification performance. Therefore, selecting an appropriate optimization technique to identify the most relevant features remains a challenging task. In this thesis, several automated deep learning-based models are proposed for diagnosing ductal carcinoma of the breast, addressing existing research gaps.

1.4 Scope of Present Work and Contributions

The primary focus of this research is to develop automated deep learning based models for the diagnosis of ductal carcinoma. Despite significant advancements in deep learning techniques, several challenges persist in cytology image analysis, including the limited availability of annotated datasets, class imbalance, staining variations, overlapping cells, and the high computational cost of deep learning models. This thesis aims to address these challenges by proposing optimized deep learning frameworks that enhance the accuracy, efficiency, and reliability of breast cancer diagnosis.

As there was a scarcity in publicly available annotated breast cytology image data, one of the primary objectives was to develop a new breast cytology image database- JUCYT, from an Indian perspective. Most researchers create their own databases, which are not publicly available and primarily focus on regions such as Europe and the USA. While a publicly available breast cytology image database (Sakia et al.) exists for the Indian context, its samples are prepared exclusively using Papanicolaou (Pap) stain. In contrast, the JUCYT dataset includes samples stained using multiple techniques, such as May-Grünwald Giemsa (MGG), Pap, and a combination of both, providing greater variability in sample representation. Additionally, most publicly available databases are annotated only at the classification level. However, in the JUCYT dataset, samples are manually annotated at both

the classification and pixel levels, enabling the segmentation of cellular structures such as nuclei and cytoplasm. Furthermore, patient details such as age and gender are recorded for future analysis while strictly adhering to ethical guidelines. Apart from dataset development, various models have been explored in this research to address key challenges. One major challenge is the limited availability of samples and class imbalance in cytology datasets. To mitigate this, generative models have been introduced to generate synthetic cytology samples with features similar to the original ones. To find the best synthetic samples from the entirely generated synthetic samples, we have introduced a selection mechanism using a fuzzy entrophy based rule. In addition to increase the volume of database, fuzzy membership functions have been explored for extracting regions of interest (ROI) from cytology images. The high-volume synthetic samples generated are utilized for data augmentation, enhancing the training dataset. The effectiveness of this augmented dataset is evaluated by analyzing the performance of Deep Convolutional Neural Networks (DCNNs) in malignancy identification. These DCNN models have been fine-tuned to further improve accuracy in classifying cytology samples. Furthermore, traditional segmentation techniques have been employed to extract crucial cellular regions from cytology images. To enhance segmentation and classification performance, ensemble methods have also been proposed. A web-based application has been developed to facilitate the automated diagnosis of breast cancer by allowing users to upload cytology images for analysis. Overall, the proposed models and the corresponding application offer significant benefits to the medical community. Additionally, due to its low-cost implementation, the system can be accessible to the general public, making advanced breast cancer diagnosis more widely available.

1.5 Thesis Organization

This thesis comprises six chapters, along with a list of references. Chapter 1 provides an overview of ductal carcinoma of the breast, its diagnostic techniques, and the importance of early detection. It then discusses cytology images, the cellular characteristics of cancerous and non-cancerous samples, and the significance of cytology imaging in breast cancer diagnosis. Additionally, this chapter reviews automated techniques for diagnosing ductal carcinoma through the analysis of

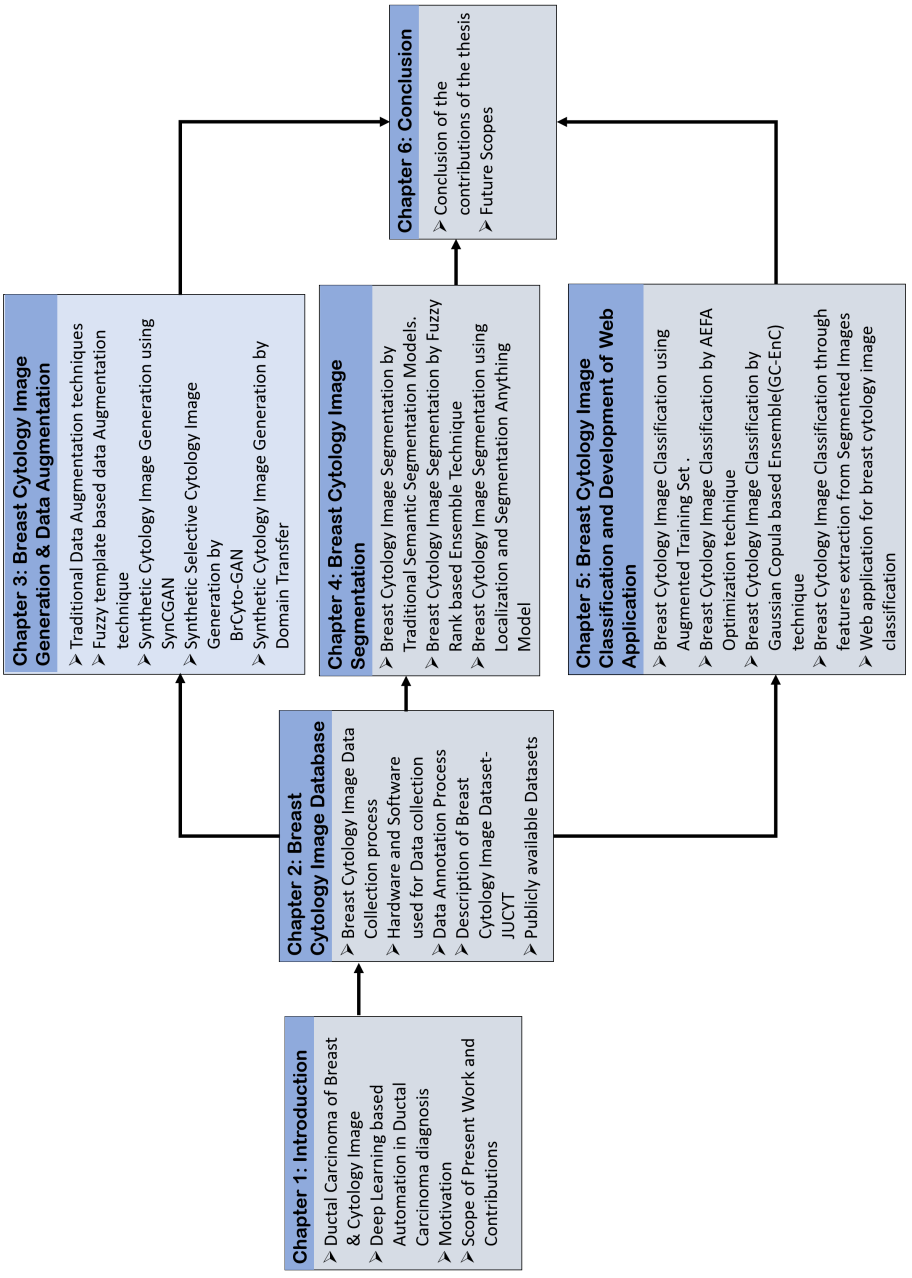


Figure 1.2: Block Diagram of Thesis Organization

digital pathology images. The role of deep learning techniques in this process is also highlighted. Finally, the chapter concludes with a discussion on the scope, objectives, contributions, and overall organization of the thesis.

Chapter 2 provides a detailed description of the newly developed breast cytology image database, covering aspects such as sample preparation, image data collection, data annotation, and preprocessing techniques. It also outlines the hardware

and software used during data collection. Additionally, this chapter includes information on publicly available breast cytology image datasets. Furthermore, details of other publicly available datasets, such as breast histopathology and cervical cytology datasets, are discussed, as they are utilized in subsequent experiments for validation purposes.

Chapter 3 discusses various data augmentation techniques for the recognition of breast cytology samples. It introduces traditional, fuzzy-based, and deep learning-based approaches for augmenting the training dataset. Novel synthetic data generation and data augmentation techniques in the field of medical image analysis is also presented. Furthermore, this chapter explores key performance metrics used to evaluate the effectiveness of the proposed synthetic data generation pipelines.

Chapter 4 presents various deep learning-based techniques for breast cytology image segmentation. It discusses traditional semantic segmentation models as well as ensemble approaches for this task. An end-to-end deep learning framework, integrating automated localization and segmentation of cellular regions is also proposed. Furthermore, this chapter reviews relevant previous works and provides a detailed analysis of the segmentation performance of the proposed models.

Chapter 5 discusses the proposed models for breast cytology image classification. It covers traditional deep learning models trained on augmented datasets and highlights an ensemble model that enhances classification performance. In addition, some morphological feature-based classifications by traditional machine learning models are examined in this study. The experimental setup, methodology, and results of the proposed models are presented in detail. Additionally, this chapter introduces a web application designed for the automated classification of ductal carcinoma.

Chapter 6 provides a comprehensive discussion on the outcomes of the thesis and highlights potential future research directions in the field of breast cytology image classification.

Finally, all the relevant references are provided at the end of the thesis.

The block diagram of over all thesis organization is mentioned in Figure 1.2.

Chapter 2

Breast Cytology Image Database

2.1 Introduction

Nowadays, breast cancer has become a very common disease among women, and the rate of mortality has been increasing in the last few years [53]. Early detection and accurate diagnosis is crucial and can improve the survival rate. When patient sees tumor or lumps on a specific organ, he or she go to the doctor and doctor suggests various diagnosis tests like CT, MRI, X-ray, biopsy etc. Among them, biopsy remains the most common and cost-effective technique [2]. There are different types of biopsy techniques, like aspiration biopsy, tissue biopsy etc. A well-known example of aspiration biopsy is fine needle aspiration cytology (FNAC) [65]. It is extensively used in the detection and diagnosis of breast lumps and also other organs like liver lesions, renal lesions, ovarian masses, soft tissue masses, pulmonary lesions, thyroid nodules, subcutaneous soft tissue mass, salivary gland, lymph nodes, and so on. FNAC plays a crucial role in the diagnosis of cancer due to its minimally invasive nature, rapid assessment capabilities and cost effective than histopathology [91]. In this test, the cellular samples are collected from breast lumps and diagnosed under microscope to distinguish their benign and malignant nature.

For the development of any deep learning based application or CAD(Computer Aided Diagnosis) based system, a high-quality and diverse dataset is crucial. The data collection process involves multiple stages like sample preparation, imaging,

annotation, and dataset structuring. However, in the field of medical imaging, data collection poses significant challenges due to high annotation costs and the need for specialized expertise.

This chapter provides a brief description of the preparation and annotation of the newly developed breast cytology image dataset (JUCYT), along with the hardware and software used for data collection. Additionally, we provide descriptions of other publicly available breast digital pathology image datasets, including both cytology and histopathology, as well as cytology datasets from other organs. Some datasets are primarily used to validate the methodologies discussed in this thesis.

2.2 Cytology Image Data Collection process

In the proposed work, we have developed a breast cytology image database, named as JUCYT. The cytology slides were prepared and images were collected at Theism Diagsis Centre, Dumdum, Kolkata, India, in the presence of professional experts.

2.2.1 Sample Collection and Slide Preparation Techniques:

Specimen Collection:

In the first step of the data collection process, the sample collection and the slide preparation are done by professional practitioners. In this process, a thin needle (normally 23-25 gauge) is injected through the skin and sufficient amount of cells are taken out for investigation [74]. This procedure was first successfully done at Maimonides Medical Center, United States, in 1981 [86]. Soon it was realized that it is relatively faster, safer, cheaper, much less painful method and trauma free diagnostic process compared to surgical biopsy. After collection, specimen is expelled into appropriately labelled glass slides with patient's unique identification. The expelled material is spread over several slides in small amounts, rather than depositing it in one large pool on a single slide to enhance the probability of error-free interpretation. This simplifies the process, in order to obtain a thin-layer preparation. Spreading of the material over the slide is usually performed by sliding another glass slide over the meterial, in order to avoid crushed artefacts and

obtain a uniform smear. Large amounts of blood are avoided to prevent clotting and fibrin trapping in the cells, which creates large cracks on the slide hindering interpretation at cellular level.

Fixation techniques and staining protocol:

After preparing the slide suitably, the fixation is required. Immediate fixation of the collected specimen is crucial; otherwise, drying artifacts may occur, potentially leading to false-positive or false-negative results in medical diagnosis. Two common fixation techniques employed in laboratories are Air-Drying and Alcohol Fixation [4]. Staining is another vital step, typically performed after fixation of cytology specimens. Most stains enhance the visibility of cellular structures by interacting with light, thereby differently illuminating specific regions of the tissue. Without staining, it becomes difficult to distinguish between different components of the sample. The intensity and distribution of illumination on a specimen depend on both the type and quantity of stain used. However, even with the same staining method, variations can arise due to factors such as the manufacturer of the stain, preservation procedures, and the condition of the specimen prior to staining.
Air Drying: After slide preparation, specimens are typically fixed by air drying—preferably within 5 minutes—to preserve cellular integrity [34]. Romanowsky stains are commonly applied to air-dried smears, with protocols including May-Grünwald Giemsa (MGG), Leishman-Giemsa (LG), and Diff-Quik stains. Among these, MGG is the most widely used technique for visualizing cell morphology and cytoplasmic details in air-dried samples. When stained with MGG, tubercle bacilli, Actinomyces, and certain fungal elements appear red, while the background takes on a pale blue hue. After application of the stain, the nucleus appears blue/black, , while the cytoplasm appears pink/orange for keratinising cells and blue/green in color for non-keratinising squamous cells.

Alcohol Fixation: Alcohol or wet fixation [34] is typically performed using either a spray fixative or by immersing slides in 95% ethyl alcohol. For alcohol-fixed slides, the Papanicolaou stain is the preferred staining method, as it offers excellent visualization of squamous cell differentiation. The staining protocol commonly includes Papanicolaou-EA-50, which targets critical regions of the nucleus and cytoplasm. Harris' hematoxylin, a key component of this protocol, is used in combination with OG-6 (Orange G) and EA-50 (Eosin Azure). OG-6 serves as a

counterstain in exfoliative cytology, particularly for smears from vaginal, cervical, and prostatic samples.

2.2.2 Image Data collection Technique:

After preparing the cytology slides, they are observed under a microscope for detailed examination. Figure 2.1 provides a concise pictorial representation of the data collection process.

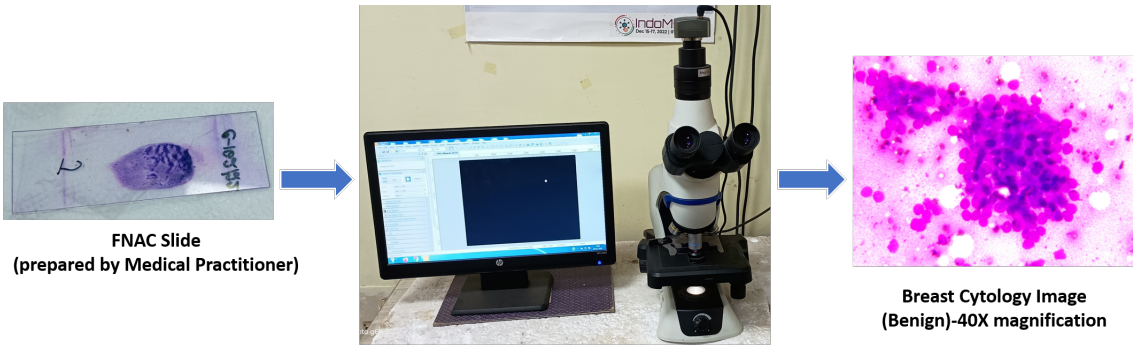


Figure 2.1: Cytology Image Data Collection Process: From slide to digital image

The microscope is equipped with lenses of varying magnifications. The observation begins with a low magnification lens, typically at 4x magnification, to locate the region of interest on the slide. Once the area is identified, the lens is progressively switched to higher magnifications, first to 10x and then to 40x, for a more detailed analysis of cellular structures. The sequence of data capture, ranging from 4x to 40x magnification of a biological slide, is depicted in Figure 2.2.

A camera module attached to the microscope captures images of the informative cellular regions identified by professional practitioners. During the image-capturing phase, critical adjustments are made to enhance the quality and characteristics of the digital cytology images. These include the following:
Focus Adjustment: Ensures sharpness and clarity of the observed region. Aperture Adjustment: Optimizes light levels and contrast for detailed visualization.

Both focus and aperture settings are modified dynamically based on the specific region being captured and the position of the slide view. This systematic process

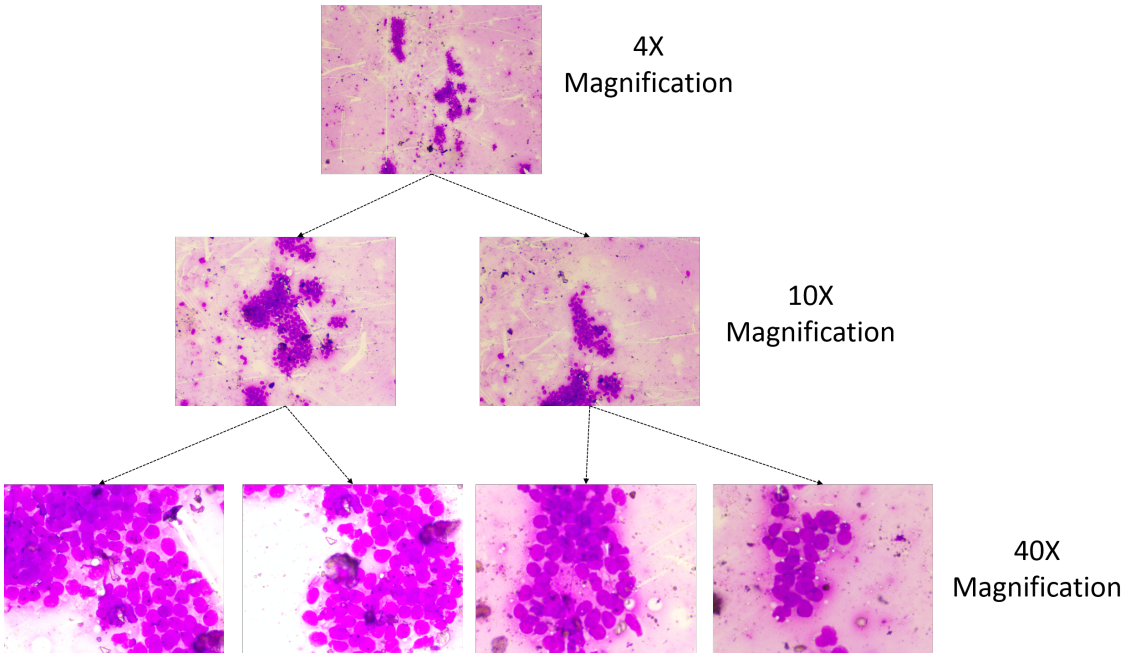


Figure 2.2: ROIs in different magnifications from Breast FNAC slide

ensures high-quality image acquisition, suitable for further analysis. From a single slide, we have captured 5-10 ROIs(region of interests) images.

2.2.3 Hardware and Software used for Data-collection

For this study, two microscopes were used: (i) a Trinocular microscope (model: Olympus CX-33) and (ii) a Binocular microscope (model: Olympus CH20i). Both microscopes were equipped with CMOS camera modules for image acquisition. Specifically, a 10 MP Magnous camera was attached to the Trinocular microscope, while a 5 MP QBC FM Scan camera was used with the Binocular microscope. Cytology images were captured using these camera-equipped microscopes. In the case of the Trinocular microscope, the third ocular tube facilitated camera attachment for image acquisition, while for the Binocular microscope, the camera was mounted directly onto the eyepiece. Imaging was carried out using two software tools: MagVision (for the Magnous camera) and QBC FM Scan, both of which are compatible with Windows and Linux operating systems. Images captured using the Magnous camera via MagVision software have a resolution of 2748×3584 pixels, whereas those captured with the QBC FM Scan camera have a resolution of 960×1280 pixels. To maintain consistency across the dataset, all images were

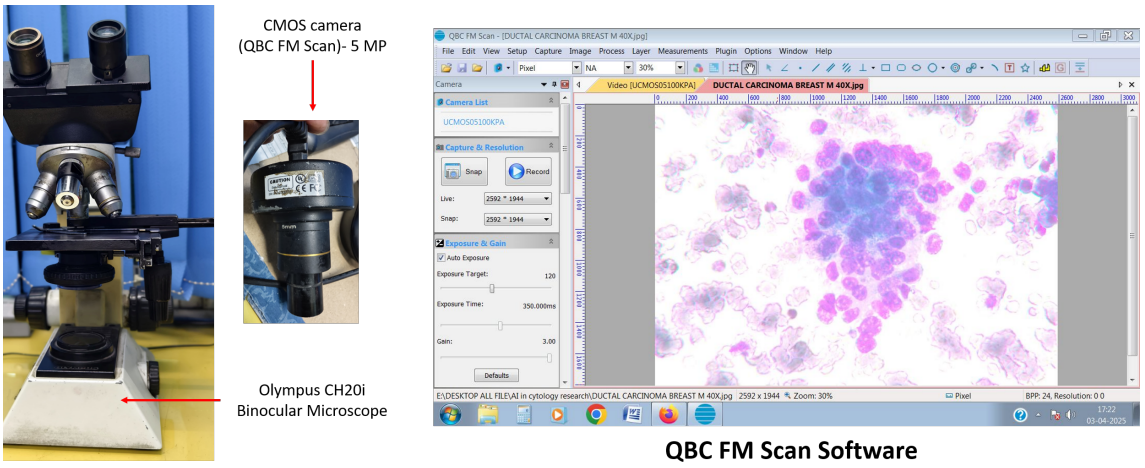


Figure 2.3: Hardware and Software used for Digital Cytology Image data collection(using Binocular Microscope)

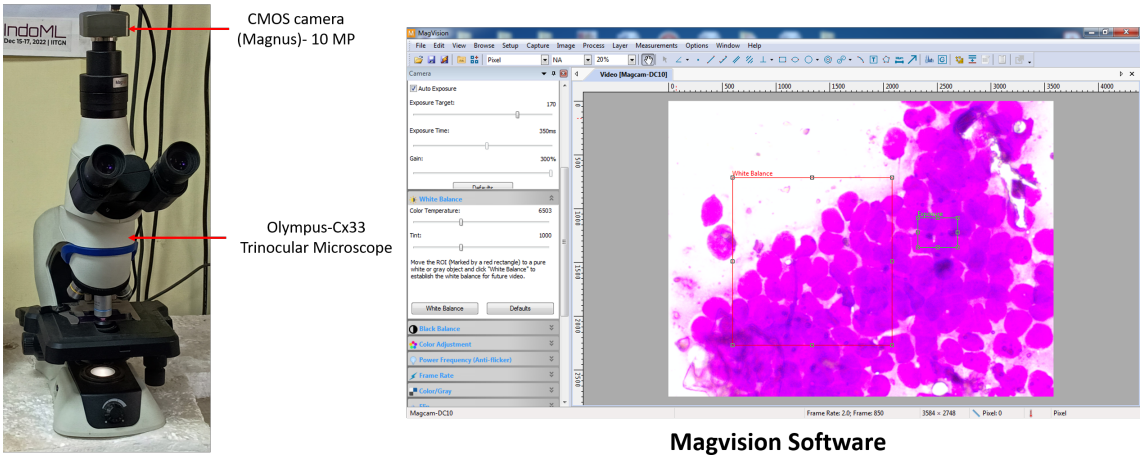


Figure 2.4: Hardware and Software used for Digital Cytology Image data collection(using Trinocular Microscope)

resized to a uniform resolution of 960×1280 pixels. This setup facilitated high-resolution image acquisition, crucial for detailed observation of the stained slides. The pictures of hardware and corresponding software are mentioned in Figure 2.3 and Figure 2.4.

2.2.4 Data Annotation Process:

After capturing the cytology images using a CMOS camera and associated software, the files were saved to the storage server's hard drive in .png format. Breast cytology images obtained from fibroadenoma or fibrocystic slides were labeled as

Benign, while those captured from ductal carcinoma slides were labeled as *Malignant*. From a single slide we captured multiple ROIs. The ROI images are then named according to slide number and slide information. If we have taken a slide(slide no: C515 & year: 2020) of ductal carcinoma of breast, and from this we have captured n ROIs in 40x magnification with label k. Then the image sample's name will be C515_Breast_2020_40x_i_k.png , where i is the ROI number(i∈[1,n]) & k be the label of slides(B for benign and M for malignant sample).

In breast cytology images, the content is divided into two regions: the background and the foreground. For segmentation and localization tasks, manual annotation of foreground objects, such as nuclei and cytoplasm, is essential. These RGB cytology images are manually labeled using Photoshop software, and the annotations are stored in .PSD format. Each cell in the images is marked layer by layer, and post-processing techniques are applied to refine the ground truth annotations. Additionally, bounding boxes for foreground objects, including nuclei and cytoplasm, is extracted from the annotated layers. The benign and malignant samples, annotated with bounding box information, are shown in Figure 2.5. Here, we evaluated the position and size of the bounding boxes using the format (x, y, w, h), where (x, y) represents the coordinates of the top-left corner of each bounding box, and w and h denote the width and height, respectively.

In the final ground truth mask, white pixels represent the informative portions (foreground), while black pixels denote the background. A flow diagram, illustrated in Figure 2.6, outlines the complete process for preparing the segmentation masks. These masks will serve as the basis for subsequent segmentation and localization tasks.

2.2.5 Image Pre-Processing Techniques:

Some traditional pre-processing techniques are applied to the raw image samples at the time of data collection to refine them. The pre-processing techniques employed here are listed below:
Image Enhancement: The contrast of raw images is improved using enhancement techniques for better visualization of cellular structures. Here some methods are employed:
Adaptive Histogram Equalization [108]: Enhances local contrast by dynamically

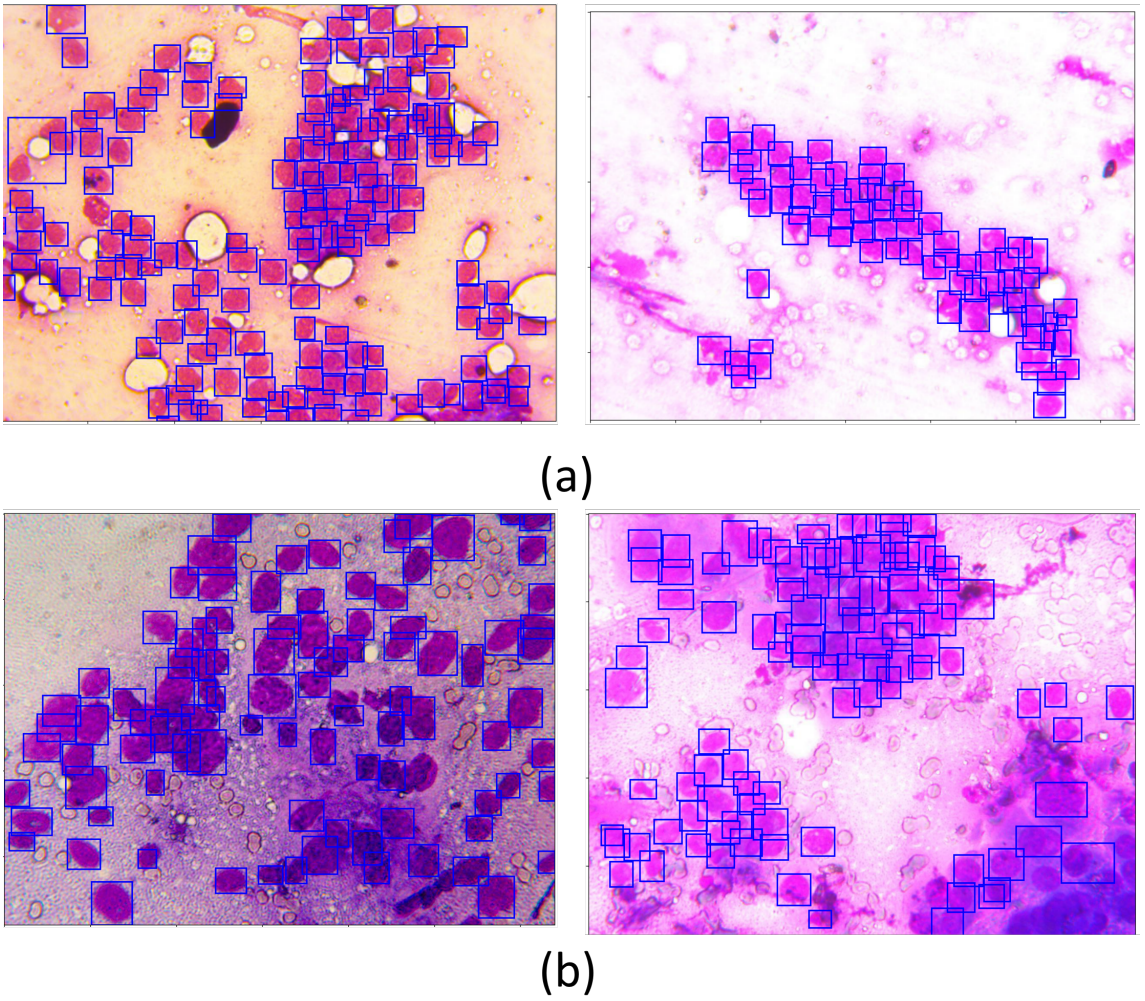


Figure 2.5: Examples of some samples with manual foreground object detections (a) Benign Sample, (b) Malignant Sample

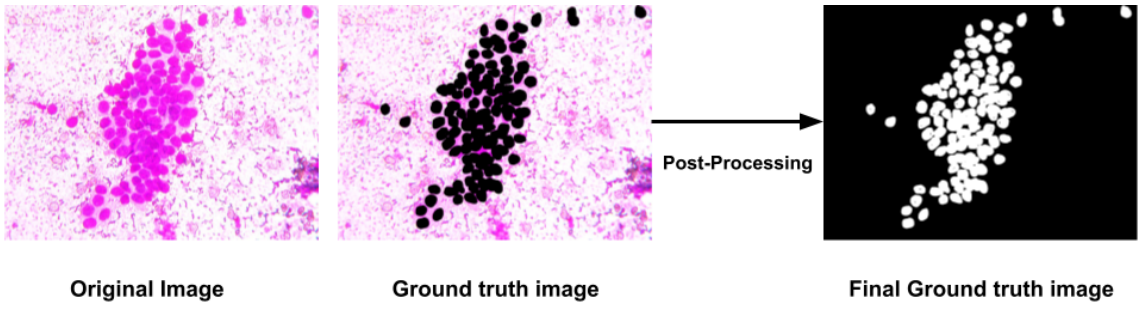


Figure 2.6: Manual Segmentation Mask Preparation for Cytology Samples

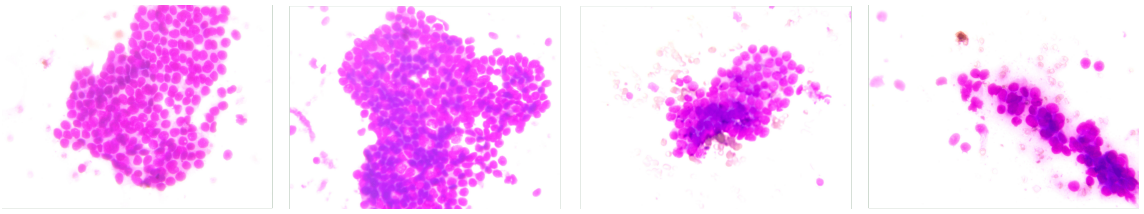


Figure 2.7: Breast Cytology Samples- Benign

adjusting brightness in different regions of the image.
Gamma Correction [37]: Adjusts brightness for consistent illumination.
Noise Reduction: Raw images often contain noise and artifacts that can hinder analysis.
Filtering techniques are employed to remove noise:
Gaussian Filter [54]: Reduces Gaussian noise while preserving image edges.
Median Filter [16]: Eliminates salt-and-pepper noise, which appears as random bright and dark spots.

These preprocessing techniques are not applied in a fixed sequence. Instead, they are selected and applied based on the condition of the captured data.

2.2.6 Description of Breast Cytology Image Dataset-JUCYT

The JUCYT database contains breast cytology image samples collected from patients of various age groups. The dataset creation process involves different steps, like, FNAC sample collection from patients, slide preparation, microscopic image data collection. We have collected total 212 cytology image samples of dimension 960×1280 pixels from the 48 patients(one slide for each patient), among which 94 are benign and 118 are malignant images. Examples of benign and malignant images at 40X magnification are shown in Figure 2.7 and Figure 2.8, respectively. The dataset is split into a standard ratio for train, validation, and test sets. The volume of the database has been steadily increasing over the years. We have utilized different subsets of the JUCYT dataset for various methodologies. The details of these subsets are mentioned in the corresponding methodology sections.

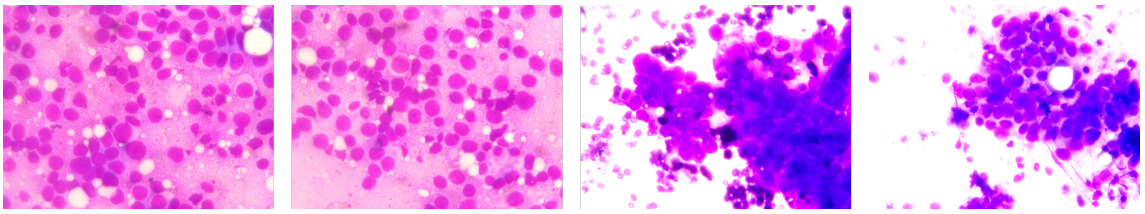


Figure 2.8: Breast Cytology samples- Malignant

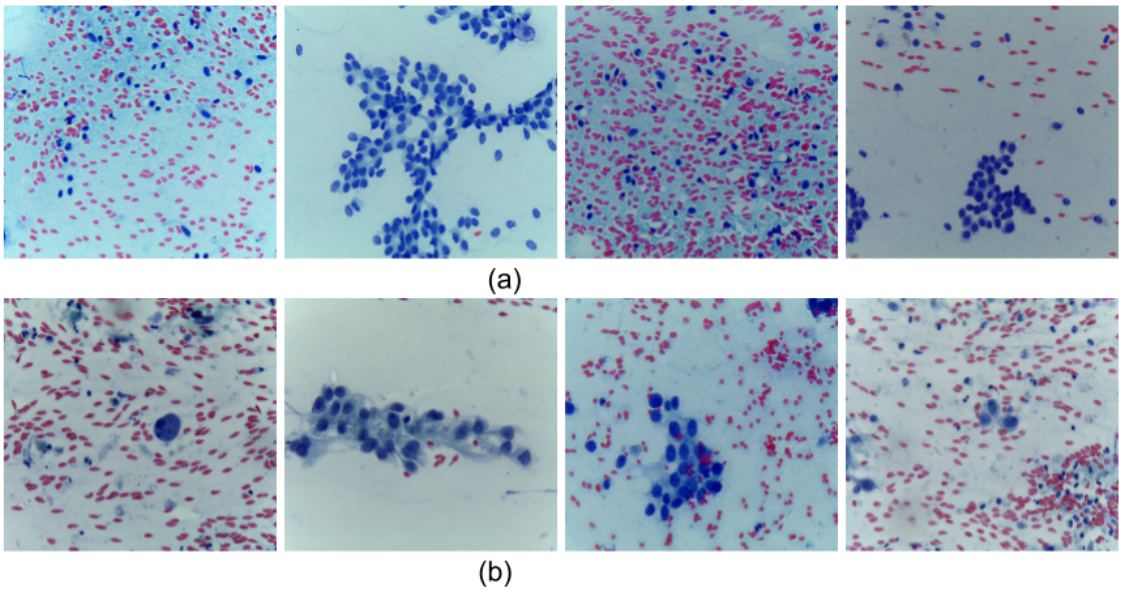


Figure 2.9: Breast Cytology Samples (Shakia et al.) (a) Benign Samples, (b) Malignant Samples

2.3 Publicly available Datasets:

Some publicly available datasets already exist in the domain of digital pathology(cytology & Histopathology) for the diagnosis of ductal carcinoma of the breast.

Saikia et al. [89] proposed a breast cytology image dataset, generated by the FNAC test. The samples were collected from Ayursundra Healthcare Pvt. Ltd, Guwahati, India. This dataset[1] consists of 212 images, which were collected from 20 patients, in which 99 are benign and 113 are malignant samples. The images are of size 2048×1536 pixels and captured by a 5-megapixel camera attached with a Leica ICC50 HD microscope. All slides were prepared by pap stain, and images are captured at 400x magnification. Examples of some breast cytology images from this dataset are mention in Figure 2.9.

[1] https://1drv.ms/u/s!Al-T6d-_/ENf6axsEbvhbEc2gUFs

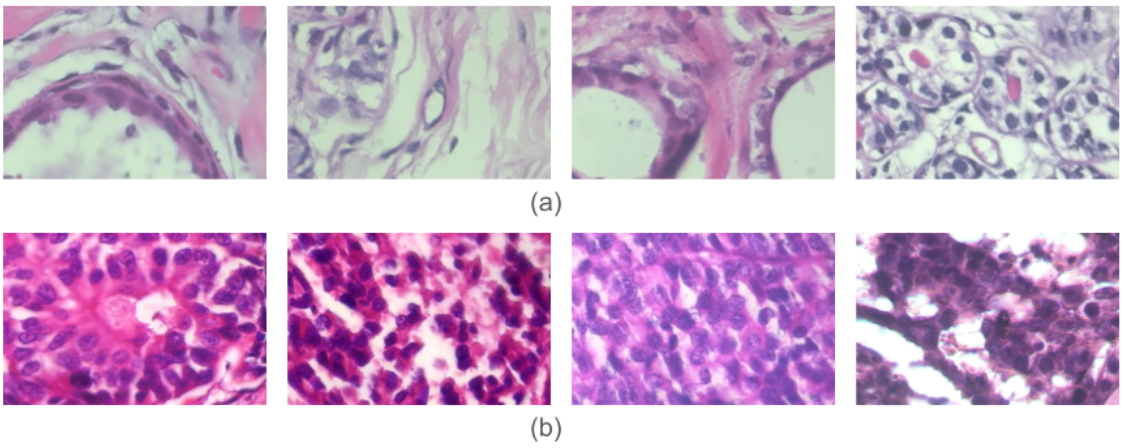


Figure 2.10: Breast Histopathology Samples(BreakHis dataset-400X Magnification)(a) Benign Samples, (b) Malignant samples

Due to a lack of publicly available breast cytology datasets and to validate the proposed models, we have explored publicly available cervical cytology datasets and investigated also the breast histopathology dataset.

BreakHis [98] a standard Histopathology image dataset[2], consists of 9109 images of breast tumor tissues, that are collected from 82 patients in different magnifications (40×, 100×, 200×, 400×). Among them, 588 Benign and 1232 Malignant samples are available at 400× magnification with 700×460 pixel resolution. The database is prepared in colaboration with the R&D Laboratory–Pathological Anatomy and Cytopathology, Parana, Brazil. Figure 2.10 presents example images from the BreakHis dataset for both benign and malignant classes.

Bioimaging 2015 breast histology (BI) dataset[3] collected from H&E stained breast histopathology sample slides. The images with resolution of 2048×1536 pixels, are captured at 200× magnification and pixel scale of 0.42 μm × 0.42 μm . The database contains 251 training images and 36 test images. It has four classes: Normal, Benign, In-Situ Carcinoma, Invasive Carcinoma where Benign and Normal classes are merged as Non-Cancerous, and In-Situ Carcinoma and Invasive Carcinoma are merged as Cancerous classes for the present work. Some examples are mentioned in Figure 2.11. After converting the binary classes, there are 143, 144 images are Non-Cancerous and Cancerous images respectively.

[2] `https://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis`
[3] `http://www.bioimaging2015.ineb.up.pt/dataset.html`

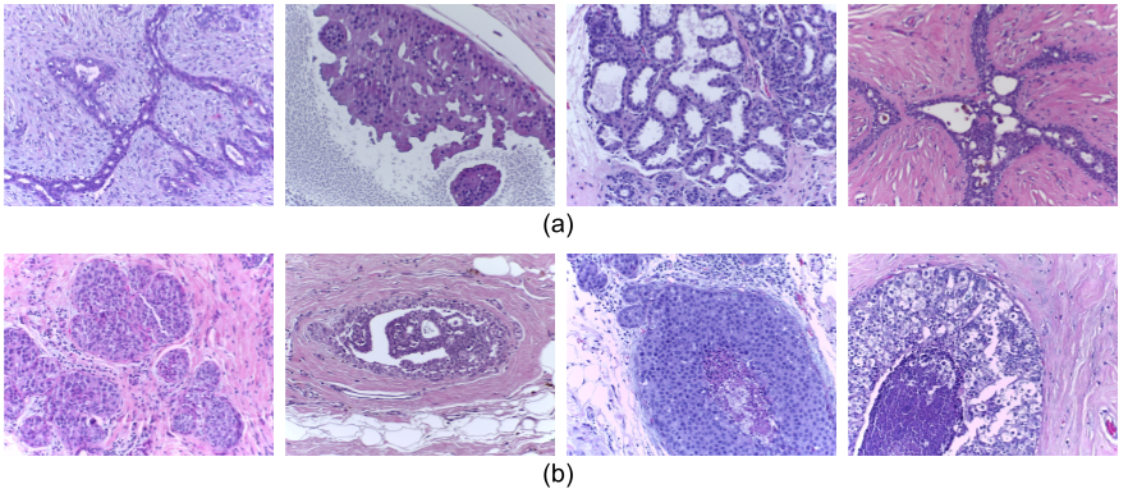


Figure 2.11: Breast Histopathology Sample(Bioimaging BI-2015 dataset)(a) Non-Cancerous Samples, (b) Cancerous Samples

SipakMeD [80] a standard cervical cytology image database[4], consisting of 966 cluster cell images with dimension 1536 $\times$ 2048 pixels. The pap smear slides are captured by CCD camera adapted to an optical microscope (OLYMPUS BX53F). The dataset has five classes, like Superficial, Parabasal, Koilocytotic, Metaplastic, Dyskeratotic. Since, we are dealing with binary class, so the abnormal cell (Koilocytotic, Dyskeratotic) classes are converted to Malignant class and the normal (Superficial–Intermediate, Parabasal) and benign cell (Metaplastic) classes are merged into Benign class. After merging, the dataset contains 505 samples under the Benign class and 461 samples under the Malignant class, respectively.

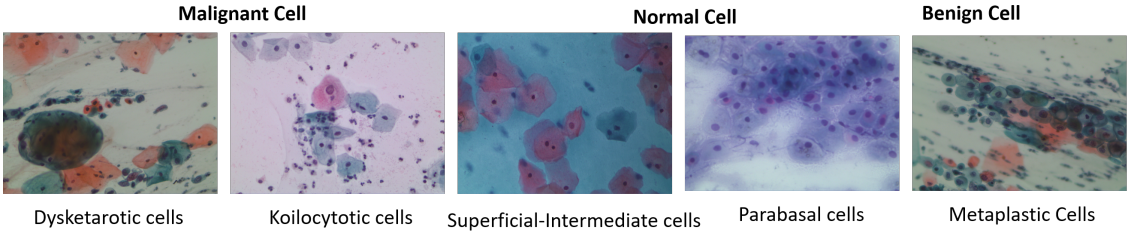


Figure 2.12: Cervical Cytology Samples (SIPaKMeD Database)

HErlev[5] is a publicly available cervical cytology image database, which consists of

[4] https://www.cs.uoi.gr/~marina/SipakMeD.html
[5] https://mde-lab.aegean.gr/index.php/downloads/

917 pap smear images (unequal image dimensions) with seven classes, like Intermediate Squamous Epithelial, Columnar Epithelial, Superficial Squamous Epithelial, Mild Squamous non-keratinizing Dysplasia, Squamous cell carcinoma in-situ intermediate, Moderate Squamous non-keratinizing Dysplasia, Severe Squamous non-keratinizing Dysplasia. The dataset consists of five different labels from the perspective of semantic segmentation. The ground truth of each image is labeled with five different colors, and these are indicative of five different portions of the Pap smear image. The red color indicates the background portions, dark blue indicates the cytoplasm portions, light blue indicates the nuclei portions, etc.

Table 2.1: Descriptions of Other Breast Cytology Image Databases

SL No.	Author	Dataset Descriptions	Dataset Source	Hardware Used
1.	Zejmo et al. [123]	500 breast cytology samples were collected from 50 patients (25 benign and 25 malignant), each with dimensions of 1583×828 pixels.	Regional Hospital in Zielona Góra, Poland.	Olympus VS120 Virtual Microscopy System.
2.	George et al. [32]	92 FNAC images(45 benign and 47 malignant samples) were captured at 10x and 40x magnifications, with a resolution of 2560×1920 pixels	Early Cancer Detection Unit, Obstetrics and Gynecology Department at Ain Shams University Hospitals	Olympus trinocular optical microscope
3.	Wolberg et al. [115]	119 images. 68 benign 51 malignant.	University of Wisconsin Clinical Sciences Center Highland Avenue Madison, Wisconsin, USA	Not Mentioned
4.	Weyn et al. [113]	83 images. 20 benign, 63 malignant	Antwerp University Hospital, Wilrijk, Belgium.	Zeiss Axiolab microscope combined with a MTI C 72CCD-camera
5.	Street [99]	569 images: 357 benign, 212 malignant.	University of Wisconsin Hospitals and Clinics	Not Mentioned
6.	Isa et al. [43]	1300 images. with four catagories like Fibroadenoma, Fibrocystic, Other benign disease, Malignant	Penang General Hospital and Hospital Universiti Malaysia, Kelantan, Malaysia.	Not Mentioned
7.	Jeleń et al. [46]	110 FNAC images. 44 , 66 images are high malignancy and intermediate malignancy respectively	Medical University of Wroclaw, Poland	Not Mentioned
8.	Malek et al. [68]	200 images in test data set. 120 benign,80 malignant	Farhat Hached Hospital, Sousse,Tunisia	Not Mentioned
9.	Filipczuk et al. [27]	275 benign and 462 malignant images were collected from 25 benign nad 42 malignant cases respectively	Regional Hospital in Zielona Góra, Poland.	Olympus VS120 Virtual Microscopy System
10.	Issac et al. [45]	334 benign proliferative diseases and 311 infiltrating duct cell carcinoma samples are collected. Images are 3264×2448 in dimension	Pathology lab of Regional Cancer Center, Thiruvananthapuram	Trinocular microscope, Leica DM2500
11.	Garud et al.[31]	37 samples. 24 benign, 13 malignant. Images are 1024×768 pixels dimension	Sub-divisional Hospital, Kharagpur, and Midnapur Medical College and Hospital, Midnapur, India	Leica DM750 microscope

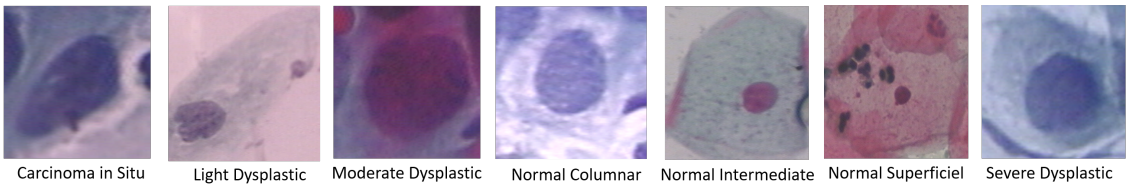


Figure 2.13: Cervical Cytology Samples (HErlev Database)

Some examples of cervical cytology images are mentioned in Figure 2.12 and Figure 2.13 for SIPaKMeD and HErlev datasets respectively.

2.3.1 Others Breast Cytology Image datasets:

Many researchers have developed breast cytology image databases for their experimental studies; however, they often provide only detailed descriptions without offering publicly accessible URLs. Some examples are briefly mentioned in Table 3.1.

2.4 Discussion

The detailed description of database, i.e. JUCYT, developed for this research has been presented in this Chapter, followed by various preprocessing techniques, and ground truths preparation for cellular objects have been discussed. In the pre-processing step, various image processing techniques are discussed to remove noise and to enhancement the contrast from the raw cytology images.

In the following chapters, various techniques for the automated diagnosis of ductal carcinoma are presented, along with implementation details.

Chapter 3

Breast Cytology Image Generation & Data Augmentation

3.1 Introduction

Data augmentation is a technique in which newly data are generated from an existing dataset to expand its size and diversity. The goal of data augmentation is to achieve better classification performance for any machine learning model. Also the deep learning-based architectures are extremely data-hungry and require a large volume of annotated high-quality data for training purposes. In the domain of medical image processing, the data collection process is both expensive and challenging due to the high annotation costs, the need for expert involvement, and the necessity of adhering to ethical guidelines. Additionally, we can face some class imbalance issues, which can hinder model performances. To mitigate this issue, data augmentation is essential.

This chapter explores various data augmentation techniques specifically designed for breast cytology images. Here, we have introduced several traditional data generation techniques and also have proposed novel synthetic data generation models. Furthermore, we conducted a study to evaluate the quality of synthetic data samples using several well-known metrics. Additionally, we implemented a two-stage synthetic image selection technique to identify high-quality synthetic samples that exhibit high feature similarity to the original samples.

3.2 Literature Survey

In the recent year, several works have been reported for data augmentation tasks in the domain of medical imaging. Most common methods for data augmentation involved affine transformations [111] such as translation, rotation, scaling, shear, flipping and so on. Also, it has been noticed that training with added noise results in a much more robust classifier. The introduction of generative adversarial networks (GANs) has brought a shift in the paradigm of generative processes in computer vision. Several approaches have made use of GANs for data augmentation. Adar et al. [29] proposed a GAN-based synthetic liver lesion ROIs generation technique. Dataset was augmented in two ways: i) the ROIs were augmented by affine transformations ii) the synthetic data was generated from ROIs using DCGAN(Deep Convolution Generative adversarial Network) and ACGAN(Auxiliary Classifier GAN). DCGAN showed better performance compared to ACGAN. Shin et al.[94] proposed a GAN based model to segment tumor in brain MRI images of two traditional datasets: ADNI and BRATs. Normal brain MRI images were segmented using an image to image translation model using CGAN[44]. Synthetic abnormal brain MRI scans were obtained from labels(tumors) by incorporating some changes in the label(e.g. increasing the size, changing the position of the label, or placing the tumor in a healthy brain MRI segmentation map). The synthetic images were used in data augmentation for training the model. Improved performance of tumor segmentation was observed by adding the synthetic data to the real data. Tom et al.[106] simulated patho-realistic ultrasound images of the IVUS dataset[1] using deep generative models. Tissue echogenicity maps were generated from the ground truth of the dataset. From these maps simulation of ultrasound images was produced using a physics-based simulator. Two-staged GAN was used to generate patho-realistic ultrasound images and stability of training state. In the first stage, images from the simulator were taken as input to GAN from which low-resolution images were generated. In the second stage, these low-resolution images from the first stage of GAN were transformed into high-resolution images. Bissoto et al.[13] suggested a GAN based model to generate high-resolution images of skin lesion of the ISIC challenge dataset[2]. Classifiers were trained on real

[1] `http://www.cvc.uab.es/IVUSchallenge2011/dataset.html`
[2] https://challenge2018.isic-archive.com/

data as well as on synthetic data. McAlpine et al. [70] introduced the StyleGAN3-adaptive discriminator augmentation technique for generating synthetic urine cytology samples. The FID score has been plateaued at 30 for this generative model. Xue et al. [118] proposed a generative model by a novel conditional GAN model. They have proposed a three-stage HistoGAN model combined with a self-attention layer after each stage of the generative model. Also, they involved a two-step selection method to identify the appropriate ones from all the generated histopathology images. In this two-step selection method they only used a single CNN architecture to extract the features. Shahidi [93] proposed a generative model, WA-SRGAN (Wide-Attention SRGAN), for super-resolution of breast histopathology images. During training, the improved Wasserstein loss with gradient penalty was applied. For the experimental study, he used the Camelyon16 and BreakHis databases, enhancing the samples with upscaling factors ranging from double to sixteen times. Man et al. [69] implemented DenseNet121-AnoGAN to classify breast histopathological images into benign and malignant categories. Instead of data augmentation, the method addressed the challenge of mislabeled patches in patch-based classification, which often arises due to errors in datasets extracted from whole slide images (WSIs). The approach integrated AnoGAN, an unsupervised anomaly detection model, to filter out mislabeled patches, and DenseNet121, a densely connected convolutional network, to extract multi-layered features from the corrected patches. The proposed method was evaluated on the BreaKHis dataset using 5-fold cross-validation, achieving a 99.13% accuracy and a 99.38% F1-score for 40X magnification images.

Many unpaired image-to-image translation models have been explored in the medical image analysis domain. Sometimes generative models are used to make multimodal data by using a single modality. There are a lot of works reported where computed tomography(CT) images are transformed into magnetic resonance(MR) images(and vice versa) by an unpaired image-to-image translation model. Welander et al. [112] proposed a study where they used CycleGan and the UNIT model for unpaired CT to MR translation. Kearney et al.[50] implemented a MR to CT image translation model to reduce the burden of the patients, who required radiation therapy. They implemented attention-aware discriminator network for cycleGAN-based synthetic data generation. It achieved structural similarity index 0.77 and PSNR 62.35. Along with the radiology domain, there are some works of image-to-image translation in the digital pathology domain. Teramoto et al.[103]

proposed a stain conversion model by using CycleGAN model. For the cytology images, the papanicolaou and giemsa stain have many complementary roles, so transferring the stains mutually is an important task in analyzing digital cytology images. Also, there exists another unpaired image translation model to generate synthetic images. Tang et al.[101] proposed AttentionGAN model for unpaired image-to-image translation task, also implemented an attention-guided discriminator to consider only the attention regions.

3.3 Data Augmentation Techniques:

This section explains several data augmentation techniques for breast cytology images. Here, we have introduced traditional data augmentation [71] along with synthetic data generation techniques.

3.3.1 Traditional data Augmentation

The primary target of data augmentation is to generate more samples from the existing training samples. There are lots of techniques like application of geometric transformations, namely translation, rotation, flipping, scaling, cropping, addition of noise, contrast enhancement, etc. which act as traditional data augmentation. Some examples of traditional data augmentations from JUCYT dataset are mentioned on Figure 3.1 and Figure 3.2 for benign nad malignant samples respectively.

3.3.2 Fuzzy template based data Augmentation technique

Extraction of random patches, or ”Region of Interest”(ROI)s from an entire image is used as one of the traditional ways for data augmentation purposes. Here, some multiple ROIs are dynamically selected from a single pathology image. This augmentation technique is applied to the breast cytology image dataset (JUCYT) and, for validation across other domains, it is also applied to breast histopathology image databases(BreakHis, BI). Instead of using other ROI selection methods, we

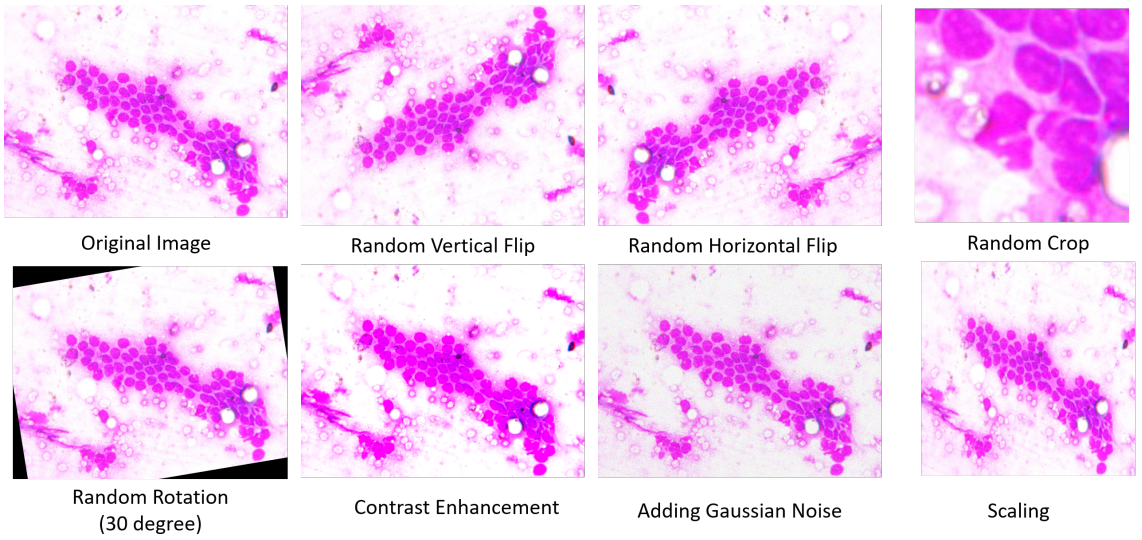


Figure 3.1: Examples of Traditional data Augmentations: Benign Case

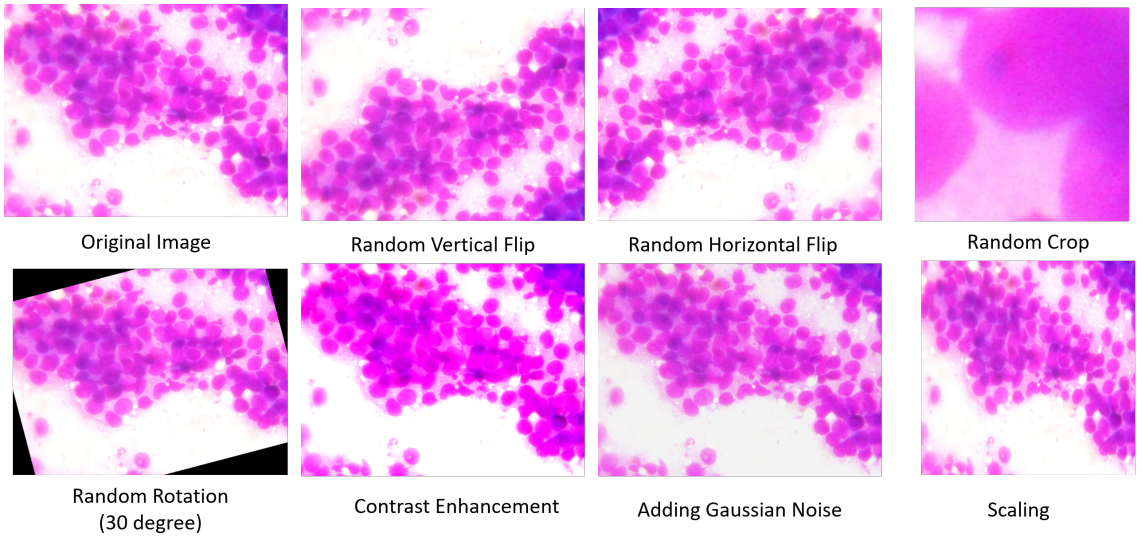


Figure 3.2: Examples of Traditional data Augmentations: Malignant Case

have used here fuzzy template based ROI detection technique to give more priority to the informative regions. In cytological images the nuclei region and its surrounding cytoplasm's carry the most informative information about the characteristics of malignant or benign tumor. Whereas for histopathology images cluster of cells or their boundaries are more informative. In the present approach we have tried to capture those ROIs as additional information which are used with the entire images to boost the performance of deep learning models. This method, mainly, consists of two major steps as discussed below:

(a) Preparation of Binary Segmentation Mask from RGB images: The first target of the fuzzy template-based data augmentation technique is to create segmentation mask from the RGB sample. Here, contrast enhancement is performed on the captured image I using the histogram stretching technique. Pixel based adaptive thresholding is done to binarize the image [58]. Finally, the Gaussian mixture clustering technique is applied to the binary mask to obtain the fine-tuned final segmentation mask. Thus, the segmentation mask I_M of same size as the image I is generated where, each background (cytoplasm) pixel value is assigned a 0, and a 1 for foreground (nucleus) pixels(The pipeline of the segmentation mask generation technique along with utilized parameters is given in Figure 3.3). Some examples of segmentation masks from RGB breast cytology samples(both benign and malignant cases) are shown in Figure 3.4.

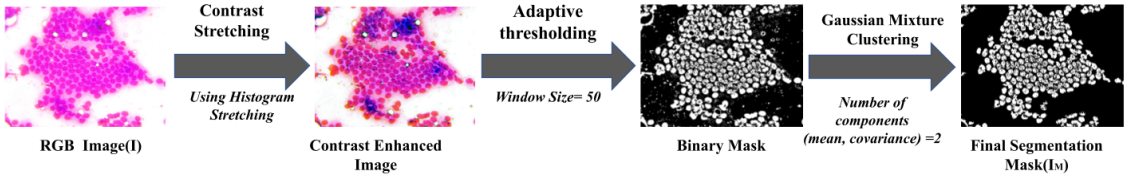


Figure 3.3: Pipeline of Segmentation Mask Generation Technique

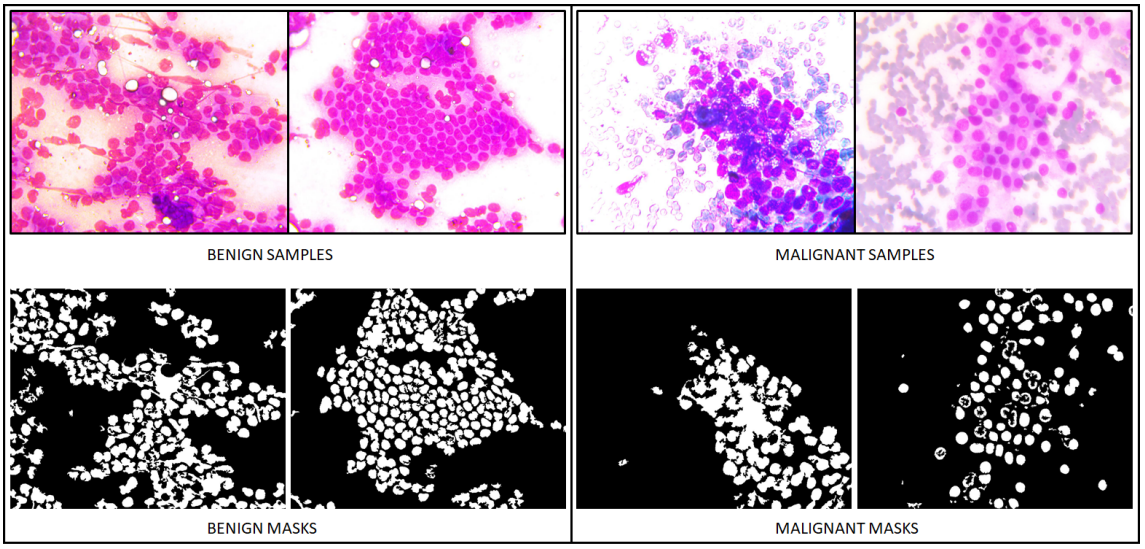


Figure 3.4: Segmentation Mask generation by [55] from real samples

(b) ROI generation by fuzzy template-based rule: After creating the segmentation mask, it is superimposed on the original image to get the masked color image, $I_S(x, y) = I_M(x, y) \times I(x, y)$ as shown in Figure 3.5. The masked color image is then converted to grayscale image I_G. Then targeted sub-images or ROIs

of sizes 256×256 are generated using the fuzzy template. For generating ROIs, intensity value at each pixel position of I_G is fuzzified to produce I_N, where fuzzy-membership value $I_N(x, y) \in [0, 1]$ is obtained using Eq. 3.1.

$$I_N(x, y) = \frac{I_G(x, y) - I_G^{\textbf{min}}}{I_G^{\textbf{max}} - I_G^{\textbf{min}}}; \forall x, y \tag{3.1}$$

$I_G^{\textbf{min}}$ and $I_G^{\textbf{max}}$ are respectively the minimum and the maximum values of intensity in I_G. Pixels in $I_N(x, y)$ having membership values in between 0.5 to 1(represented as $\mu(x, y) \in (0.5, 1]$) are selected to form a descending ordered set $I_{NS}(x, y)$. The maximum membership value is then taken from the set for generation of ROI. For equal membership valued pixels, centrally located pixel (x_c, y_c) is preferred. Five square windows (centre, right, left, bottom and top) of dimensions $R \times C$ (256×256 for our experiment) are then fitted on I_N based on the location of the selected pixel (x_c, y_c) as shown in Figure 3.6. Where, $x_c, y_c \geq \delta$; $x_c \leq (R - \delta)$; $y_c \leq (C - \delta)$ and δ=$R/2$=$C/2$ i.e., half of the size of window. Satisfying the above conditions five windows centering (x_c, y_c), $(x_c, y_c + \delta)$,$(x_c, y_c - \delta)$,$(x_c + \delta, y)$, and $(x_c - \delta, y_c)$ are constructed. The number of pixels N_{PC} having $\mu(x, y) \in (0.5, 1]$ are calculated within a region. A region is selected as an ROI if, $N_{PC} \geq 0.30 \times R \times C$. High membership values (0.5-1) are used for covering the important foreground portions, i.e. nuclei and its surroundings which are primarily responsible for distinguishing benign from malignant. ROIs are constructed from regions having at least 30% significant foreground materials. From a single coordinate five different regions (centre, left, right, top, bottom) are considered for selection as ROIs if the above criterion is met. Thus, we have iteratively generated a set of ROIs from a single image given by, $I_{ROI} = [ROI_i]_{i=1}^{k}$, where k is the total number of ROIs of image I. If $I_{ROI} = \{\emptyset\}$, the whole image is considered as an ROI and resized to 256×256. Figure 3.7 portrays the method for generation of ROIs from intensity values of a sub-image.

To confirm the selection of an ROI_i in the final set, we determine the intersection of two ROIs as:

$$f(ROI_i) = \arg\max_j \{\ ROI_i \cap ROI_j\ \} < 0.8 \tag{3.2}$$

where, $\forall i, j; i \neq j$ and $ROI_i, ROI_j \in I_{ROI}$.

Those ROI_is are considered that satisfy Eq. 3.2. ROI_is are eliminated in this process whose $f(ROI_i) \geq 0.8$. The process is summarized in Algorithm 1. Thus,

Algorithm 1: Selection of Fuzzy based ROI

Input:$I(x,y)$; **Output:**I_{ROI}
Fuzzy ($I(x,y)$)
$I_N(x,y)$ computed from $I(x,y)$ by Eqn3.1
Generate $\mu(x,y)$ from the $I_N(x,y) \in (0.5,1])$
$I_{NS}(x,y) \leftarrow PUSH(\{(x,y)|\mu(x,y)\})$
Sort $I_{NS}(x,y)$ in descending order
Initialize k=1
for p=1 to $|I_{NS}(x,y)|$ **do**
    M← membership value of $I_{NS}(x,y)(p)$
    $I_{NS}^{POP}(x,y) \leftarrow POP(\{(x,y)|\forall(x,y) = M\})$
    $Temp_(x,y) \leftarrow$ CentrallyLocated (I_{NS}^{POP})
    $I_{NS}(x,y) \leftarrow PUSH(\{(x,y)|\forall(x,y \in I_{NS}^{POP}$ except CentrallyLocated (I_{NS}^{POP})
    Fit windows $W_{(R\times C)}^{i}$ where $i = 1,..,5$ on I_N w.r.t $Temp_{(x,y)}$ as described in Section **??**
    for $i = 1$ to 5 **do**
        $N_{PC} \leftarrow |\{\ (x,y)|W_{(R\times C)}^{i}(x,y) > 0.5\ \}|$
        if $N_{PC} \geq 0.30 \times R \times C$ **then**
            $ROI_k = W_{(R\times C)}^{i}$ and PUSH ROI_k into I_{ROI}
        end if
        k ← k + 1
    end for
end for
if $I_{ROI} == NULL$ **then**
    return $I(x,y)$
end if
for $i,j = 1$ to k **do**
    Select ROI_i and ROI_j from I_{ROI} such that $i \neq j$
    if *f(*ROI_i*) ≥* 0.8 *// (using Eq. 3.2)* **then**
        I_{ROI} ←I_{ROI}- ROI_i // Eliminate ROI_i
    end if
end for
return I_{ROI}

from a single image multiple ROIs are generated. Image segment corresponding to each selected ROI is cropped from the original image and considered as a new sample of its base class. These generated ROIs are used for next level classification task.

Some examples of ROIs(dimension 256 ×256) which are generated from the original breast cytology and histopathology images by fuzzy template based rule, are shown in Figure 3.8.

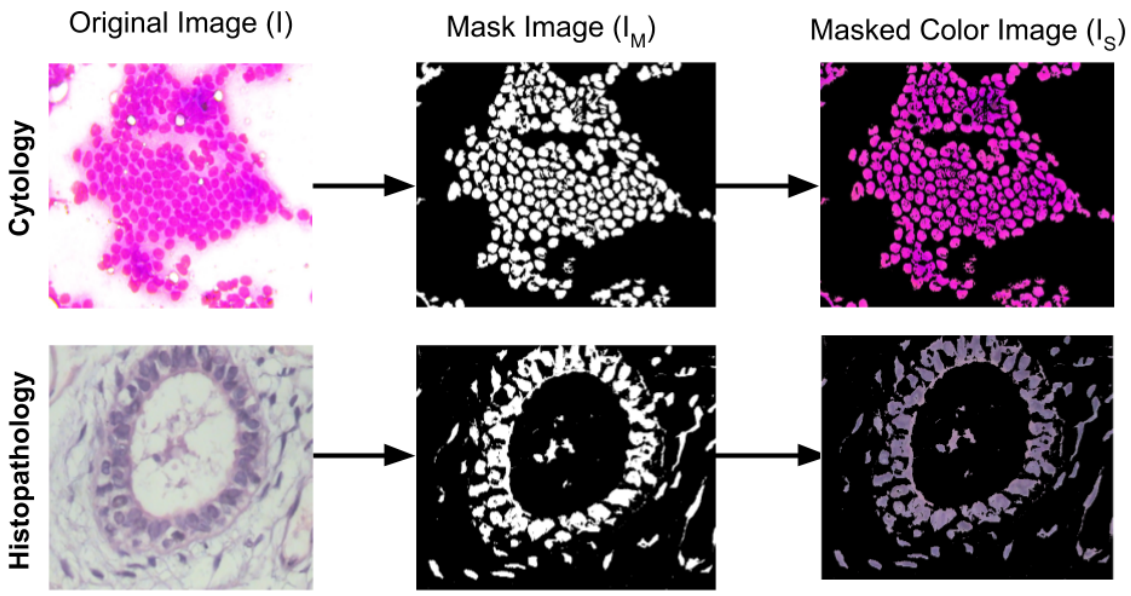


Figure 3.5: Example of masked images generated from input images

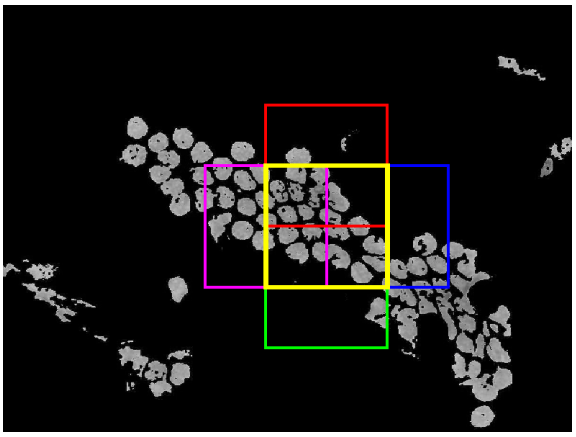


Figure 3.6: ROIs generated on a Fuzzified Image are indicated by different colours, i.e. Yellow for center, Red for Top, Blue for right, Pink for left and Green for bottom ROIs respectively.

3.3.3 Synthetic Cytology Image Generation using SynCGAN

Synthetic sample generation is a well-known data augmentation technique for enhancing classification performance across various domains. In this work, a hybrid generative model-SynCGAN is proposed specifically for the breast cytology domain.

Data augmentation pipeline

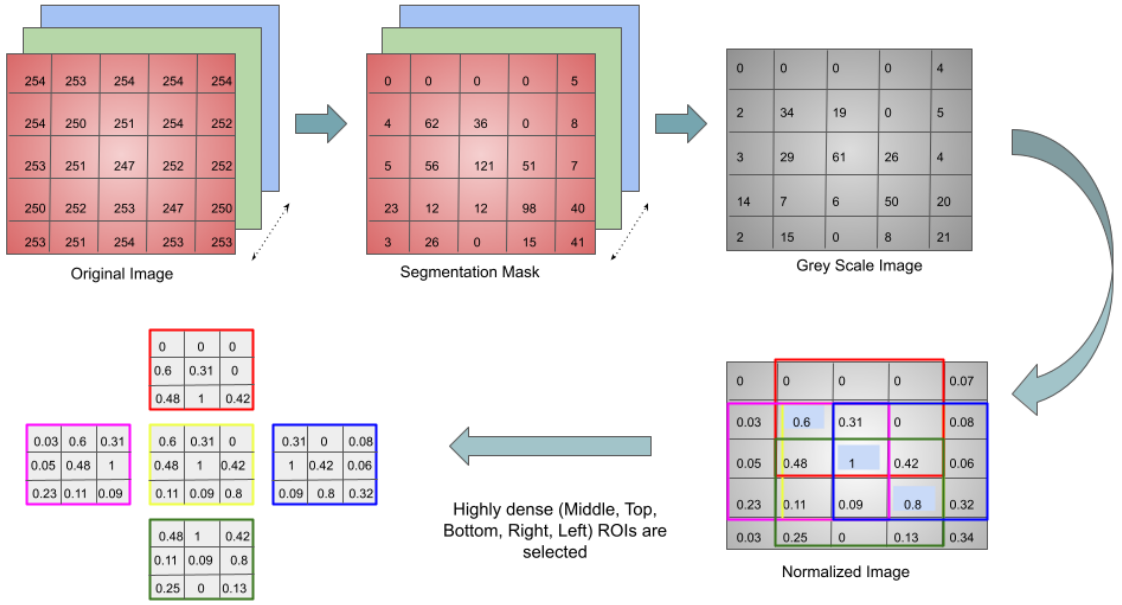


Figure 3.7: Automated ROIs extraction process from the intensity values of a sub-image

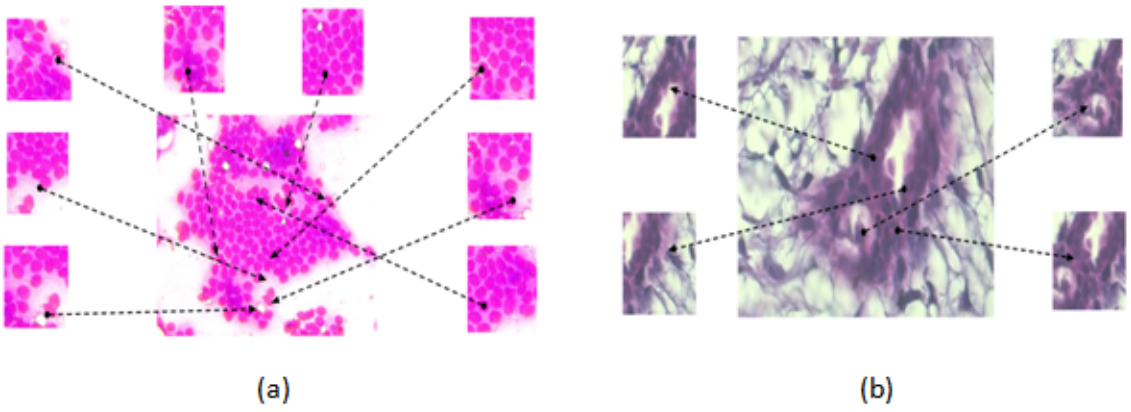


Figure 3.8: ROIs (256×256) generation from original image (a) Cytology (b) Histopathology

The proposed data augmentation pipeline consists of three steps. The block diagram of the SynCGAN model is shown in Figure 3.9. Firstly, segmentation masks are collected from real images using an unsupervised method. Secondly, a CGAN is trained on these pairs of images and auto-generated segmentation masks. Two sets of synthetic segmentation masks are generated using a GAN for each of the two classes. Finally, these synthetic segmentation masks are used to guide the previously trained CGANs to generate patho-realistic samples for data augmentation.

Step I: Mask generation

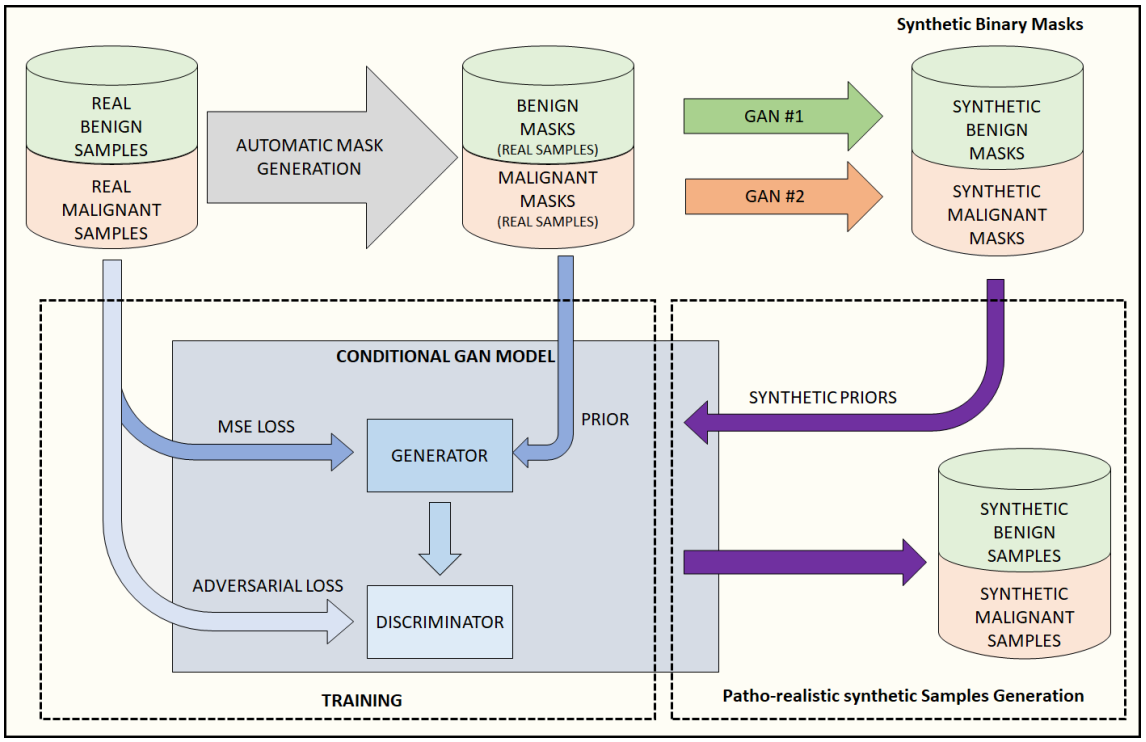


Figure 3.9: SynCGAN: The proposed data augmentation pipeline

The proposed methodology requires a set of pixel-level annotations to guide a CGAN for data generation. Due to the lack of hand-annotated samples, an unsupervised approach was used for nuclei segmentation. Here, the previously mentioned(see Section 3.3.2(a)) binary segmentation mask preparation technique [55] is applied for that task. The presence of an unsupervised mask generation technique alleviates the necessity of a large amount of training data.

Step II: RGB Synthetic Sample Generation by CGAN

Generating RGB images from scratch using a traditional GAN is difficult as the generation procedure can be represented as a prediction of 256-dimensional multinomial distribution across three channels for each pixel. However, the segmentation mask is simply a pixel specific binomial distribution which is much easier to predict when starting without a predefined prior. Thus a CGAN [72] must be trained which takes the segmentation masks generated in the previous step as a prior and a generator loss is reduced against the corresponding RGB image. We derive inspiration from the pix2pix network [44]. For the generator, we use a modified UNet like architecture. Normally the UNet architecture uses transposed convolution for upscaling the feature maps. However, that results in checkerboard artifacts due

to overlap of kernels during the fractional stride. Instead, bilinear interpolation is opted for upscaling the feature map followed by a 3×3 convolution layer for refinement. The discriminator network has been directly implemented using the PatchGAN discriminator as demonstrated in [44]. The discriminator attempts to detect real and fake samples from the dataset and the generator respectively. The objective function V can be written as:

$$\min_G \max_D V(D,G) = \mathbb{E}_{(x,y)\sim p_{data}(x,y)}[\log(D(x|y))] + \mathbb{E}_{y\sim p_{data}(y)}[\log(1 - D(G(y)))]. \tag{3.3}$$

Here G and D refers to the generator and discriminator. x represents the RGB sample, y represents the corresponding binary segmentation mask which serves as a prior for the generator. The x and y samples are drawn from the input data distribution $p_{data}(x, y)$ that consists of RGB images and their segmentation masks in the current scenario.

At every iteration, the discriminator and the generator are trained alternatively as was performed previously. During the training of the discriminator, the segmentation mask and its corresponding RGB image are concatenated on its channel dimension. It is then passed through the discriminator (Patch GAN)[44] and discriminator loss of the real image is calculated as below

$$D_{loss} = -\log(D(x)) + \log(1 - D(G(y))), \tag{3.4}$$

where x represent samples from the input database and y refers to segmentation masks of those samples. A binary cross-entropy loss function is used to calculate the adversarial loss. While training the generator, the segmentation masks are passed through the generator network and the loss is calculated. The loss has two components denoting the adversarial loss exhibited by the discriminator and the mean square error between the generated sample and the actual RGB image from the dataset that corresponds to the mask y.

$$G_{loss} = -\log(D(G(y))) + \lambda\, \text{MSE}(G(y), x|y) \tag{3.5}$$

where λ is the weight of Mean Squared Error (MSE) loss. The weight of the mean squared error loss is set to 100 based on empirical analysis on a small validation set. x and y represent samples from the RGB image and the segmentation mask dataset.

Step III: Synthetic Segmentation Mask preparation by GAN

A conditional GAN(CGAN) usually generates synthetic samples conditioned by some predefined priors. In the current scenario, the CGAN has been trained to generate samples from segmentation masks highlighting the spatial distribution of nuclei across the cytoplasm. To generate patho-realistic synthetic samples a class-specific prior distribution is necessary. For that purpose, we train a GAN model [35] to generate binary segmentation masks based on a randomly drawn seed from a gaussian distribution. The most straight forward method to generate synthetic samples would have been to train an end-to-end GAN. However, it has been noticed in ablation studies that without a prior the quality of outputs is very poor. The primary reason being the complexity of predicting the intensity value of a pixel. Given the output image has three channels, each pixel exists within a search space with 256^3. However, a binary segmentation mask is a much easier output to predict given that each pixel belongs to a binomial distribution. On the other hand, the shape information encoded within these segmentation masks is quite informative about the class of the samples, namely, benign or malignant. We train two separate GANs trained on segmentation masks belonging to each of the predefined classes.

The objective of the GAN network is simply defined as

$$\begin{aligned}\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}[\log(D(x))] \\ +\mathbb{E}_{z\sim p(z)}[\log(1-D(G(z)))]\end{aligned} \tag{3.6}$$

Here x refers to samples drawn from the input data distribution $p_{data}(x)$. z refers to randomly sampled priors from a Gaussian distribution $p(z)$. G and D refers to the generator and discriminator network. The architecture of the GAN used in the current work is very similar to the one described in the previous step. It consists of a generator inspired from UNet whose transposed convolutions have been replaced with bilinear interpolation for upscaling and a convolution layer

for refinement. The discriminator is derived from the PatchGAN discriminator as demonstrated in [44]. During the training phase, the discriminator loss is given by,

$$D_{Loss} = -(\log(D(x)) + \log(1 - D(G(z)))) \tag{3.7}$$

and the generator loss is given by:

$$G_{Loss} = -\log(D(G(z))) \tag{3.8}$$

However, due to very low number of samples, the discriminator was too overpowering and saturates at a very early stage. To deal with this issue some additional measures were taken as described below:

- **Label smoothing:** The labels for real and fake samples are set as 1 and 0 by default. To enforce some fuzziness in the system, a random number between 0.9 and 1 was taken for real samples and a random number between 0.1 and 0 was taken for fake samples while training. However, this is unnecessary while training the generator as we want to bottleneck the learning curve of the discriminator and not the generator.
- **Randomly flipping labels:** To even further confuse the discriminator at some random iterations real samples are labeled as fake and vice versa. This confusion results provide some breathing space for the generator so that it can learn the requisite features.

Step IV: Generation of patho-realistic synthetic images using best Generators

The final step generates class specific patho-realistic synthetic samples. According to the pipeline discussed in section 3.3.3, at first the class-specific GANs are used to generated segmentation masks(see Step III). Then the generated segmentation masks are fed into the trained CGAN model(see Step II) to obtain the RGB synthetic images. The synthetic data distribution is modeled as,

$$p_{data(z)} = \mathbb{E}_{z \sim p(z)}[G_{CGAN}(G_{GAN}(z))] \tag{3.9}$$

where, G_{CGAN} and G_{GAN} refers to the generators of the CGAN and GAN model described in Step II and III respectively.

While a simple CGAN trained on a prior denoting the class of samples could also be viable, having a richer representation like a mask provides more flexibility in terms of variations in the synthetic dataset.

Examples of synthetic masks generated by the GAN model for both benign and malignant cases are shown in the top row of Figure 3.10. The corresponding patho-realistic, RGB-colored breast cytology samples produced by the proposed model are displayed in the bottom row of the same figure.

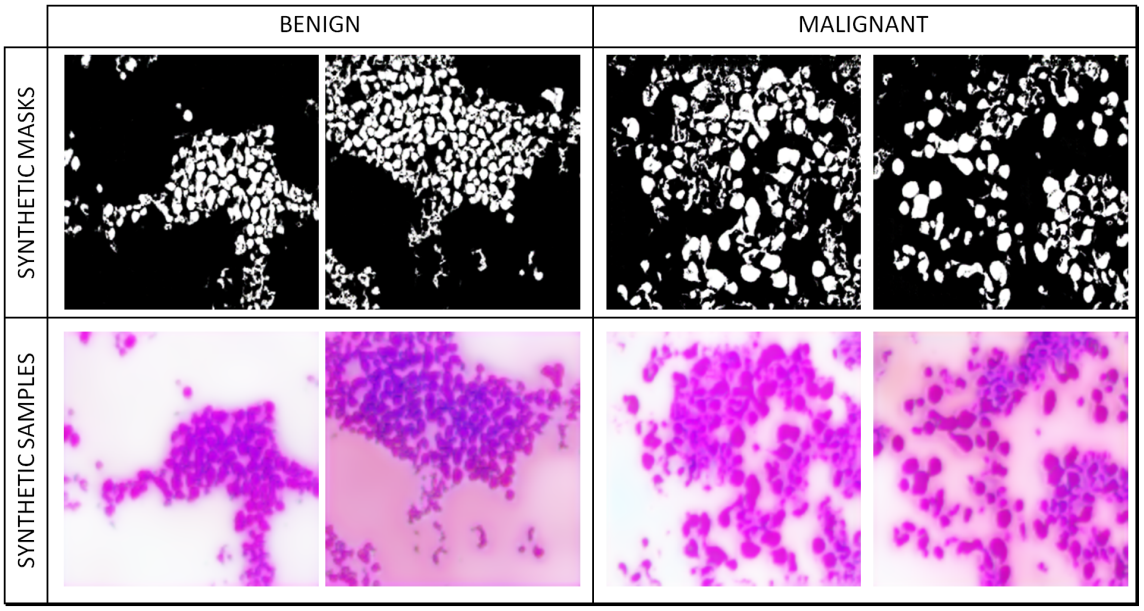


Figure 3.10: Synthetic masks generated by class specific GANs (top) are fed as priors to the conditional GAN to generate patho-realistic samples (bottom) for benign(left) and malignant(right) classes separately

3.3.4 Synthetic Selective Cytology Image Generation by BrCyto-GAN

In the previous SynCGAN model, synthetic samples with dimensions of 256×256 pixels were generated. However, these synthetic samples were generated without proper constraints. Our primary objective was not only to generate synthetic data but also to produce high-quality, high-resolution patho-realistic samples, which have close features similarity to the original training data. In medical imaging, particularly for pathology images, high-resolution image generation is very crucial to visualize the nuclei, cytoplasm, and other important objects, which will be helpful for the diagnosis of carcinoma. Here, we introduce the Super-Resolution GAN

(SRGAN) model in conjunction with the previously described SynCGAN model to achieve this objective. Also here we have implemented a fuzzy ensemble-based two-stage synthetic image selection technique to distinguish the synthetic samples which have close features to original training samples.

The proposed selective image-generative model(BrCyto-GAN) consists of five distinct steps.

Step I: The masks are prepared by an unsupervised segmentation approach.

Step II: The synthetic binary segmentation masks of dimension 256 $\times$ 256 are generated using the traditional GAN model from noise.

Step III: A super resolution mask of dimension 512 $\times$ 512 is generated using SR-GAN model from synthetic binary mask of dimension 256 $\times$ 256 which is the output of the Step II.

Step IV: The high-resolution masks are colored by the CGAN model.

Step V: A two-stage selection process is proposed to select the synthetic images, which have closer features to the original cytology images. Class-specific probabilistic scores have been calculated for the synthetic samples by first considering each CNN model separately. Then, a fuzzy entropy based ensemble rule is used to combine these probability scores to make the final decision on the selection of each synthetic image.

The block diagram of proposed synthetic data generation technique is shown in Figure 3.11. Also, the training process of proposed Hierarchical GAN model is briefly mentioned in Figure 3.12. The detailed scheme is described briefly in the following subsections.

Preparation of Segmentation Mask

Preparation of segmentation mask digital pathology images is a complex task, due to the presence of various overlapped objects like nuclei, cytoplasm, etc. Also, in any supervised learning based segmentation approach, we need a large number of manually labeled segmentation masks which is time-intensive and costly. So, in this work, we have used an unsupervised approach to segment the RGB-colored image (I). The previously described(See Section 3.3.2(a)) unsupervised

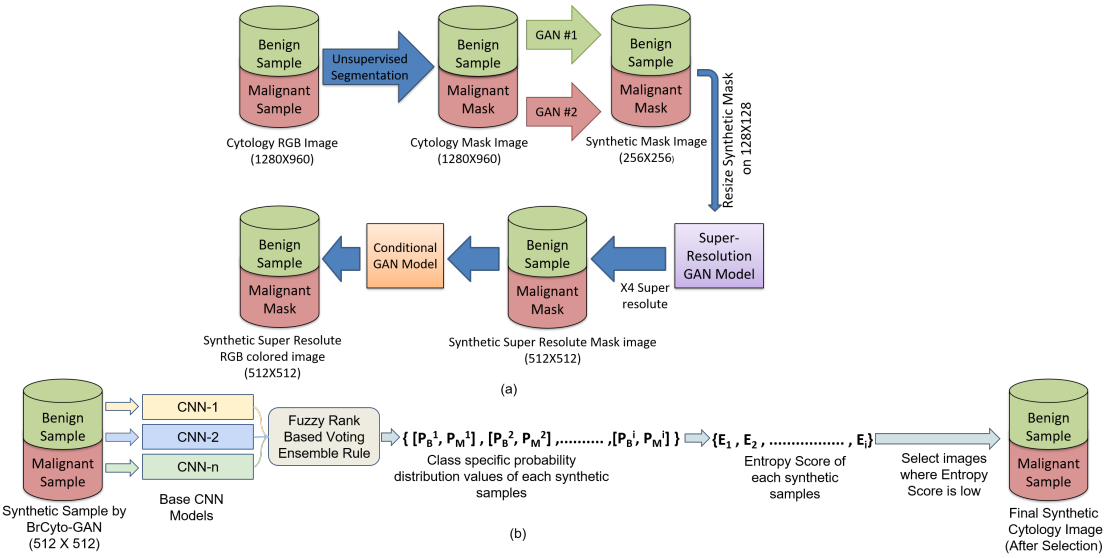


Figure 3.11: Block Diagram of the Proposed Methodology showing;(a) Synthetic Image Generation, (b) Ensemble based Selection strategy

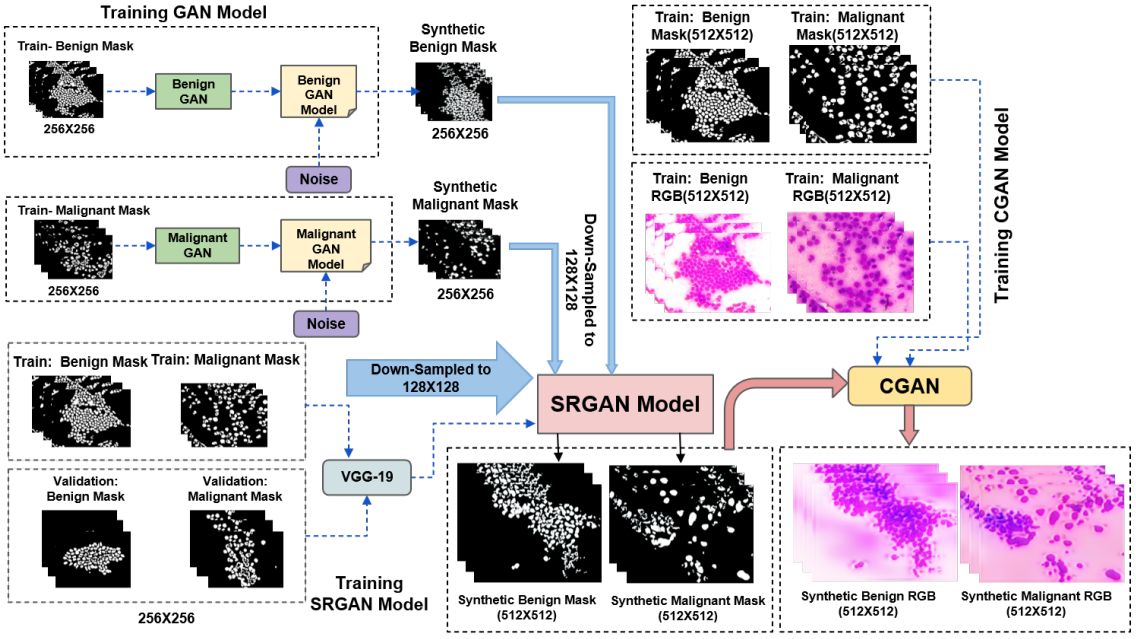


Figure 3.12: Block Diagram showing the training process of the proposed Hierarchical GAN model

segmentation technique(S_{seg}) have also been used for that case. So, the binary segmentation mask(I_M) is prepared from the original image I, i.e. $I_M \leftarrow S_{seg}(I)$.

Synthetic Segmentation Mask Generation: Generation of RGB cytology image from noise is difficult due to complex color variations within the image. Experimentally, it is also found that it leads to unsatisfactory results. Figures 3.13

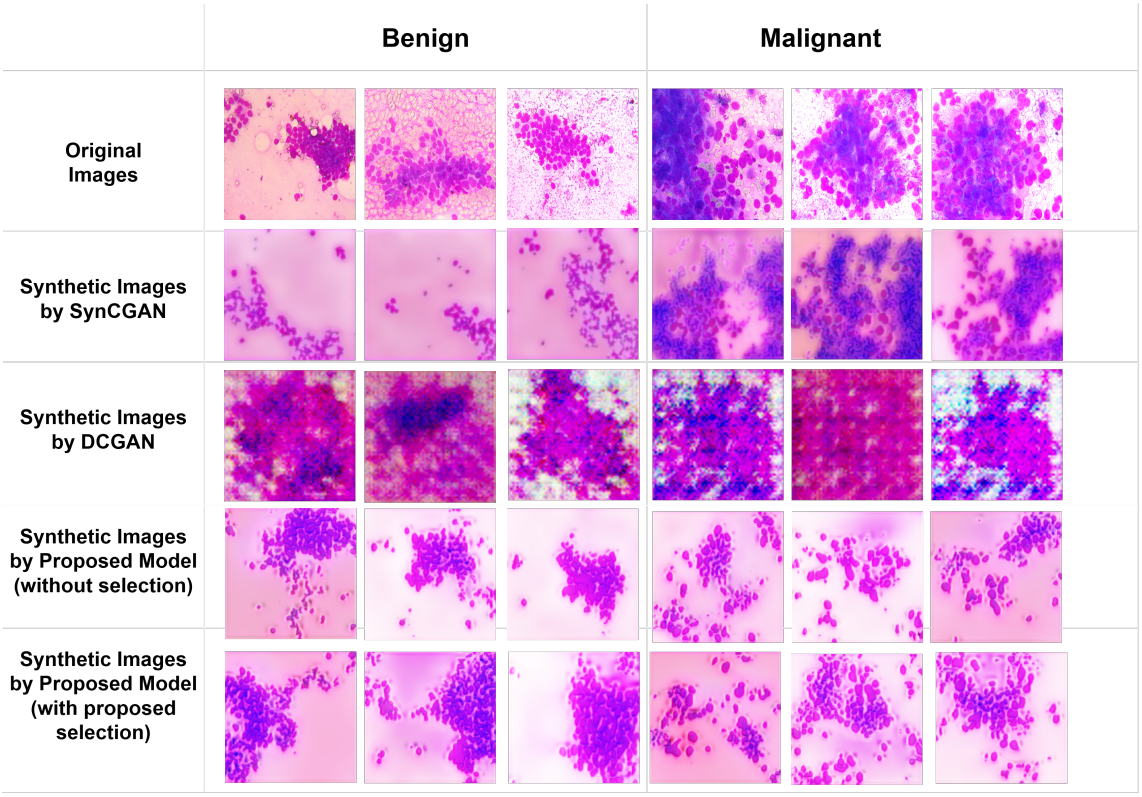


Figure 3.13: Examples of Original and Synthetic Breast Cytology image(JUCYT) samples by different generative models

and 3.14 illustrate that the synthetic RGB images generated from noise by DCGAN model contain more dissimilar features than the original one. Moreover, we have less amount of data which makes it even more challenging to generate high-resolution images from scratch. Hence, we have first generated the synthetic binary masks to guide this generation workflow. Also, generating binary images using a traditional GAN is much more efficient than trying to generate RGB-colored images because it is easier to capture a binomial distribution. So, as the second step of BrCyto-GAN model, the synthetic binary masks are prepared by the same methodology as the previously mentioned(see Section 3.3.3:Step III) SynCGAN model. The architecture of the generator, discriminator model and the corresponding training process is also mentioned in that section.

The segmentation mask I_M is used for this step of the GAN based generation. The generator loss and the discriminator loss are as given by,

$G_{loss} = -\log[D(G(z))]$

$D_{loss} = -[\log[D(I_M)] + \log[1 - D(G(z))]]$

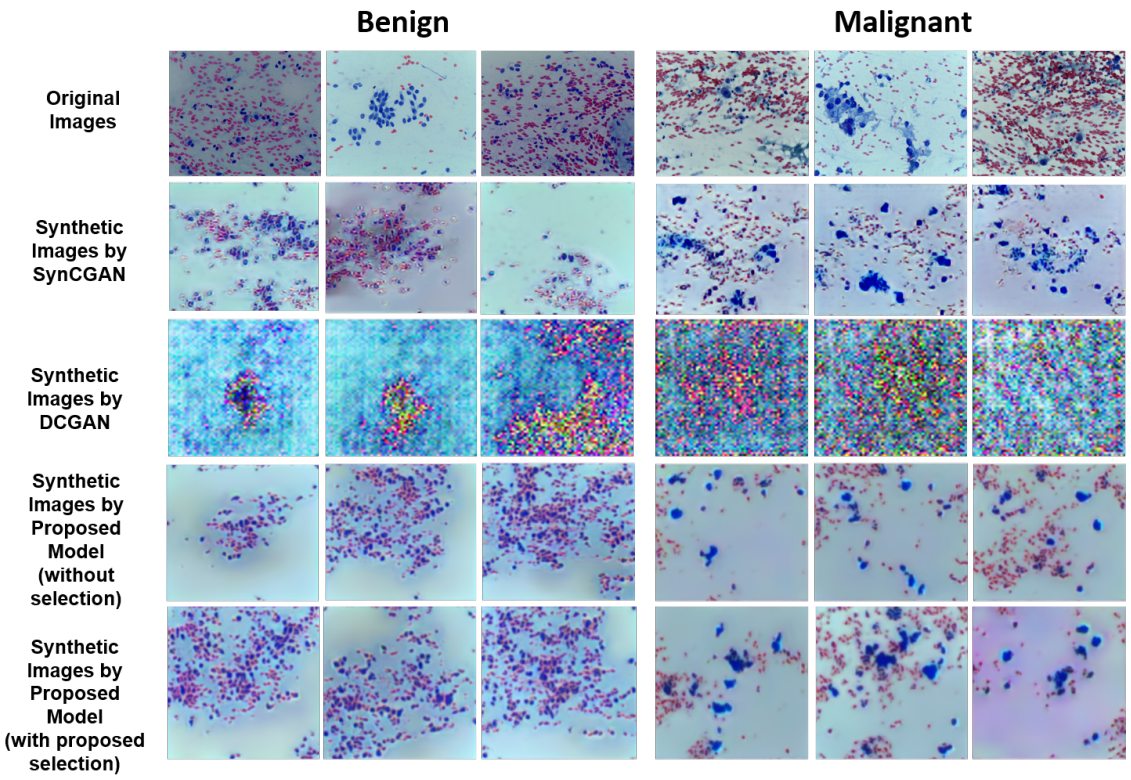


Figure 3.14: Examples of Original and Synthetic Breast Cytology image (Saikia et al.) samples by different Generative models

where, z refers to random noise, which is sampled from the Gaussian distribution with mean=0 and standard deviation=1. G and D represent generative and discriminator models, respectively.

Synthetic High Resolution Mask Generation

In this step, we have generated high-resolution (512×512 dimensional) synthetic masks by SRGAN model from the synthetic segmentation masks of dimension 256×256 generated from the previous stage. The digital pathology image at super-resolution provides enhanced structural details of crucial cell elements such as nuclei and cytoplasm. Furthermore, it has the potential to boost the performance of diagnostic models [96]. Similar to GAN architectures, the Super Resolution GAN (SRGAN) constitutes two integral parts; Generator and Discriminator. The generator produces some super-resolution data, based on the probability distribution, while the discriminator tries to guess whether data is coming from the input dataset or the generator. The system receives input of low-resolution synthetic mask images and generates a higher-resolution version of those images as output.

In the generator module, we have used SRResNet model [109], which has 16 identical residual blocks. In the discriminator, we have used the architecture similar to that mentioned in Ledig et al. work[63]. It contains eight convolutional layers with an increasing number of 3×3 filter kernels. The segmentation masks(i.e. I_M) obtained from the first step(i.e. Section 3.3.4), are used further to train the SRGAN model in the present work. SRGAN uses perpetual loss function (LSR) which is the weighted sum of two loss components: Content loss and Adversarial loss. The Content loss calculates the loss between the generated image and the real image based on the ReLU activation layers of the pre-trained VGG-19 network. The Adversarial loss is the loss function that forces the generator to produce the images more similar to high-resolution images by using a discriminator that is trained to differentiate between high-resolution and super-resolution images. The discriminator and the generator are trained simultaneously in every iteration like traditional GAN model.

The unsupervised mask I_M is first down-sampled into 128×128 and passed through the SRGAN model. Let I_M^{LR} and I_M^{HR} be the segmentation masks of dimension 128×128(Low-resolution mask), and 512×512(high-resolution mask) respectively.

The adversial min-max problem is defined as:

$$min_{\theta_G} max_{\theta_D} \mathbb{E}_{I_M^{HR}}[\log D_{\theta_D}(I_M^{HR})] + \mathbb{E}_{I_M^{LR}}[\log(1 - D_{\theta_D}(G_{\theta_G}(I_M^{LR})))]$$

$LSR = Loss_{VGG}^{SR} + 10^{-3} * Loss_{Gen}^{SR}$ where, $Loss_{VGG}^{SR}$ and $Loss_{Gen}^{SR}$ are called the content loss and adversarial loss respectively.

$Loss_{VGG}^{SR} = \sum\sum |\phi(I_M^{HR}) - \phi(G(I_M))|$

$Loss_{Gen}^{SR} = \sum -\log D(G(I_M))$

Patho-realistic RGB Colored Image Generation: This is the fourth and final step of our RGB image generation procedure. In this step, we used a CGAN model to accomplish this task. CGAN is a type of GAN that involves the conditional generation of images by a generator model. In the current scenario, the CGAN is trained to generate samples from segmentation masks highlighting the spatial distribution of cellular objects like nuclei and cytoplasm.
In this step, we have used modified UNet-like architecture, which is inspired by the traditional pix2pix model. Traditionally the UNet architecture uses transposed

convolution to up-sample the feature maps. But, in that case, checkerboard artifacts may manifest on the image due to the overlap of kernels during the fractional stride. For implementing this CGAN model, the PatchGAN discriminator [17] has been used as a discriminator network. Similar to the earlier phase, the discriminator network is employed here to differentiate between real and fake samples. The segmentation mask generated in *Step I*, is used as a learnable class-specific prior to guide the CGAN model to generate this pathorealistic sample [22]. At every iteration, the discriminator and the generator are trained alternatively as was performed previously. During the training of the discriminator, the segmentation mask and its corresponding RGB image are concatenated on its channel dimension. After that, it is passed through the PatchGAN discriminator model, and the discriminator loss of the real image is calculated as follows:

$$G_{loss} = = \log(D(G(I_M^{HR}))) + \lambda * MSE(G(I_M^{HR}), I|I_M^{HR}) \tag{3.10}$$

$$D_{loss} = = \log(D(I)) + \log(1 - D(G(I_M^{HR}))) \tag{3.11}$$

where I represents samples from the input database and I_M^{HR} refers to super resolution segmentationmasks of those samples. A binary cross-entropy loss function is used to calculate the adversarial loss. While training the generator, the segmentation masks are passed through the generator network and the loss is calculated. The loss has two components denoting the adversarial loss exhibited by the discriminator and the mean square error between the generated sample and the actual RGB image from the dataset corresponding to the mask.

Selection of High-Resolution RGB Image:

In the first step of the proposed generative model, an infinite number of noise vectors from a Gaussian distribution have been used to generateinfinite number of synthetic masks. After sequentially applying SRGAN and CGAN to these masks, we can generate an infinite number of RGB images, but there is no guarantee that all of those image samples have the feature similarity to the original RGB samples. Also, it will be impossible to select patho-realistic synthetic images from a huge number of synthetic images by using visual recognition. The training dataset is enhanced with randomly chosen synthetic images; however, these selections might

not always share feature similarities with the original images, potentially reducing the performance of classification. So, we develop an image selection model that uses an ensemble rule based on fuzzy logic to eliminate synthetic samples that are more dissimilar to the original images. The proposed selection model consists of two phases. Firstly the class-specific probability distribution values of each synthetic cytology image are extracted by traditional CNN models like DenseNet-161($\mathbb{D}-161$), InceptionNet-V3($\mathbb{I}-3$) and ResNet-18($\mathbb{R}-18$), which are trained on the original cytology image dataset.

Let, B_i^j and M_i^j be the class-specific probability distribution values of i^th synthetic image with respect to CNN model j of Benign and Malignant class respectively, where $B_i^j + M_i^j = 1$ and $j = 1, 2, 3$. Now the fuzzy ranks $R_B^{i_1}$, $R_B^{i_2}$ for Benign class are calculated where, $R_B^{i_1} = 1 - tanh(\frac{(B_i^j-1)^2)}{2}$, $R_B^{i_2} = 1 - exp(-\frac{(B_i^j-1)^2)}{2}$. After that, the fused rank score RS_B^j for the Benign class is calculated as $RS_B^j = R_B^{i_1} \times R_B^{i_2}$. Similarly, the fuzzy ranks($R_M^{i_1}$, $R_M^{i_2}$) and fused rank score(RS_M^j) for malignant class are calculated. After that, the fused score for the Benign class(FS_B) is calculated as $FS_B^i = \sum_{i=1}^{\mathrm{N}} RS_B^i$(where N= No of CNNs are used for ensemble). Also, the fused score for the malignant class (FS_M^i) is calculated in a similar way. So, the fusion score-based tuple of i^{th} synthetic data is in the form $[FS_B^i, FS_M^i]$. This class-specific fusion score of the i^{th} synthetic image is predicted by the minimum value of class label fused score. To represent the fused scores as probability distribution values, first they are inverted (i.e. $[1 - FS_B^i, 1 - FS_M^i]$) and then passed through softmax function ($SOFTMAX(x) = \frac{e^x}{\sum e^x}$). After this ensemble rule, the predicted class probability of i^{th} synthetic image is more uncertain. To select only the synthetic samples with a certain probability distribution, we have used entropy ranking in the next step of the selection process. So, the final class-specific distribution value of i^{th} synthetic data is $[P_B^i, P_M^i]$. The flow diagram of the ensemble technique is mentioned in Figure 3.15. In the next step, the entropy score of each synthetic data(E_i) is calculated $E_i = -(P_B^i \log P_B^i + P_M^i \log P_M^i)$. After that, the median of entropy scores with respect to a specific class is calculated. Then the final synthetic images will be selected where the entropy score will be lower than the median entropy score of that class. The fuzzy Entropy based selection process is summerized in Algorithm 3. The low entropy score indicates the selective synthetic samples are more certain concerning the classes and the samples with more certain labels give better performances in classification tasks.

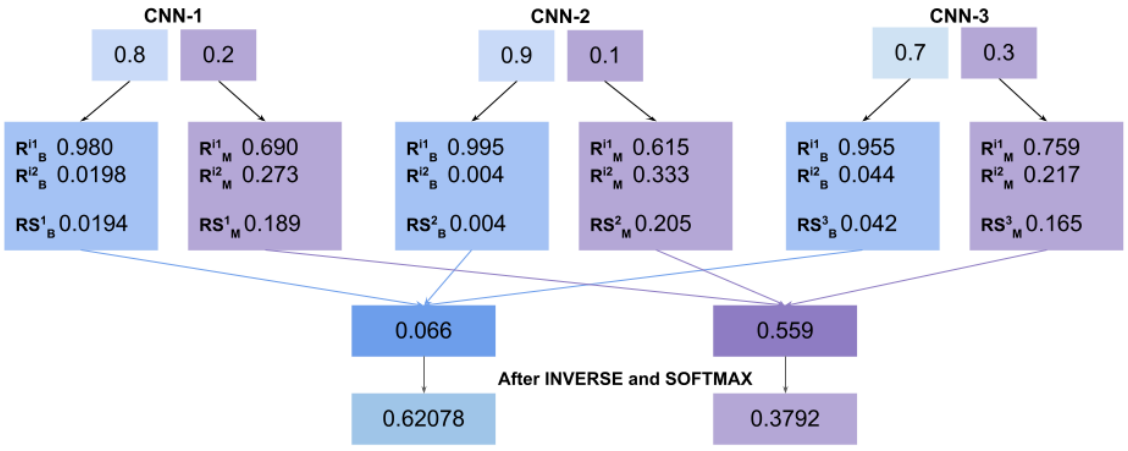


Figure 3.15: Illustration of Ensemble process(CNN 1: DenseNet-161, CNN 2: ResNet-18, CNN 3: InceptionNet-V3)

Examples of some selective synthetic samples(both Benign and Malignant cases), generated by the proposed BrCyto-GAN model, are shown in the last row of Figures 3.13(for JUCYT) and 3.14(for pap stain breast cytology).

Algorithm 2: Proposed Fuzzy Entropy-based Selection Technique

Input: $[B_i^j M_i^j]; j = 1, 2, 3$
Output : D_C
$R_C^{i_1} = 1 - tanh(\frac{(C_i^j - 1)^2)}{2}$, $R_C^{i_2} = 1 - exp(-\frac{(C_i^j - 1)^2)}{2}$.
Calculated for $\forall C$ = B and M, for each j
$RS_C^j = R_C^{i_1} \times R_C^{i_2}$; Calculated for $\forall C$ = B and M, for each j
$FS_C^i = \sum_{k=1}^{|j|} RS_C^k$;$\forall C$ = B and M
$[P_B^i, P_M^i] \leftarrow [SOFTMAX(1 - FS_B^i), SOFTMAX(1 - FS_M^i)]$
$E_i = -(P_B^i log P_B^i + P_M^i log P_M^i)$
$D_C = \{\ \}\ \forall C$
for i←1 to N_C **do**
    $E_C \leftarrow MEDIAN(E_i)$

    if $E_i < E_C$ **then**
        $D_C = E_C \cup i$
    end if
end for

Experimental Setup for SynCGAN Model: The JUCYT FNAC cytology image dataset was used for this study. All images have a resolution of 960×1280 pixels and were captured at 40X magnification (see Chapter 2). A total of 156 cytology images were analyzed, comprising 77 benign samples and 79 malignant samples. The dataset was randomly divided in the ratio of 3:1:1 into training, validation, and testing set with an equal number of samples in each class. The ratio of synthetic data to original training data was maintained at 2:1. Thus 90 synthetic

Table 3.1: Number of samples in the original and synthetic dataset

Dataset	Training	Validation	Testing
Original Images	90	30	30
Synthetic Images	180	-	-

training images were generated per class. These synthetic samples are selected by visual interpretation and domain knowledge. The data distribution in details are mentioned in Table 3.1.

The CGAN model trained for a maximum of 200 epochs and the best model was saved based on minimum generator loss on the validation dataset. While the GAN model was trained for almost until the generator loss saturated(1600 epochs). For both the cases adam optimizer was used. All the experiments were conducted on Nvidia GTX 1060 GPU.

Experimental Setup for BrCyto-GAN model

In this proposed work, we have used two breast cytology image datasets:(a) JUCYT ,(b) Breast Cytology Pap stain dataset(Saikia et al.[89]). Data distribution for each of three sets considering a binary class is given in detail in Table 3.2 for both the datasets. It also describes the data distribution of augmented training sets after various selection processes, including the proposed selection using a single CNN and multiple CNNs, traditional data augmentation techniques, and other ensemble-based selection methods. The classification performance of each selection during training is detailed in the next chapter.

The training process for generating synthetic samples are conducted using an NVIDIA GTX Geforce 970 GPU with Intel core-i5 and 16 GB RAM.

Metrics for the quality evaluation of Synthetic Samples:

Several image quality metrics like IS(Inception Score) [11], FID (Fréchet inception distance) [121], KID (Kernel Inception distance) [107] are genrally used to evaluate the quality of synthetic samples. The details are mentioned in the following:

KID Score:

$$KID = MMD(f_{real}, f_{fake}) \tag{3.12}$$

Table 3.2: Detailed dataset distribution of different training rules

Description	JUCYT		Saikia et al.	
	Benign	**Malignant**	**Benign**	**Malignant**
Train set	77	79	60	68
Test set	9	20	20	23
Validation set	8	19	19	22
Total Synthetic data	4000	4000	4000	4000
Selected Synthetic data (Fuzzy Ensemble of $\mathbb{R}-18$,$\mathbb{I}-3$,$\mathbb{D}-161$+ Entropy Rule)	2077	2079	2060	2068
Selected Synthetic data (Fuzzy Ensemble of $\mathbb{R}-18$, $\mathbb{I}-3$+Entropy Rule)	2077	2079	2060	2068
Selected Synthetic data (Fuzzy Ensemble of $\mathbb{R}-18$, $\mathbb{D}-161$+Entropy Rule)	2077	2079	2060	2068
Selected Synthetic data (Fuzzy Ensemble of $\mathbb{I}-3$,$\mathbb{D}-161$+Entropy Rule)	1710	2079	2060	2068
Selected Synthetic data (Only Fuzzy Ensemble)	849	429	1316	1681
Selected Synthetic data (Selection by InceptionNet-v3)	1670	1678	1182	2068
Selected Synthetic data (Average Probability + Entropy)	2077	2079	2060	2068
Traditional Data Augmentation	539	553	420	476

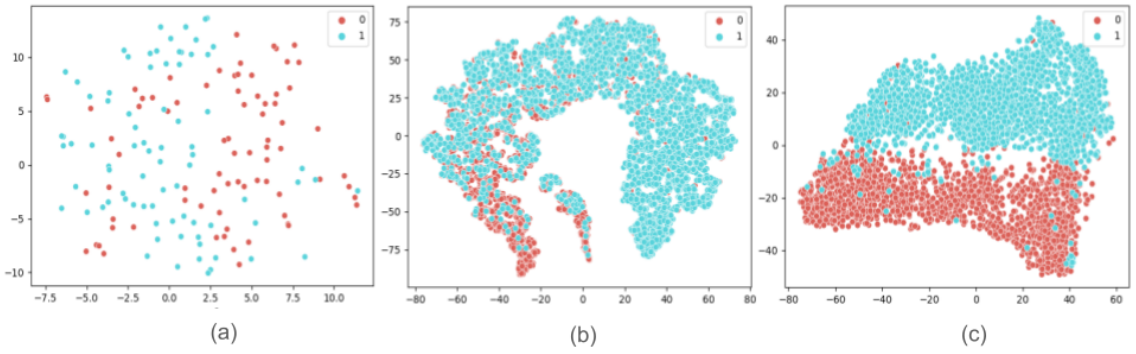


Figure 3.16: t-SNE plot of JUCYT dataset (a) Original Train Set, (b) Augmented Train Set (before selection) (c) Augmented Train Set (after selection; ensemble by three base classifiers)

where, *MMD* is the maximum mean discrepancy [8] and f_{real}, f_{fake} are represented by the extracted features by the InceptionNet-v3 model(which is pretrained on ImageNet) from the real and synthetic samples respectively.

IS Score:

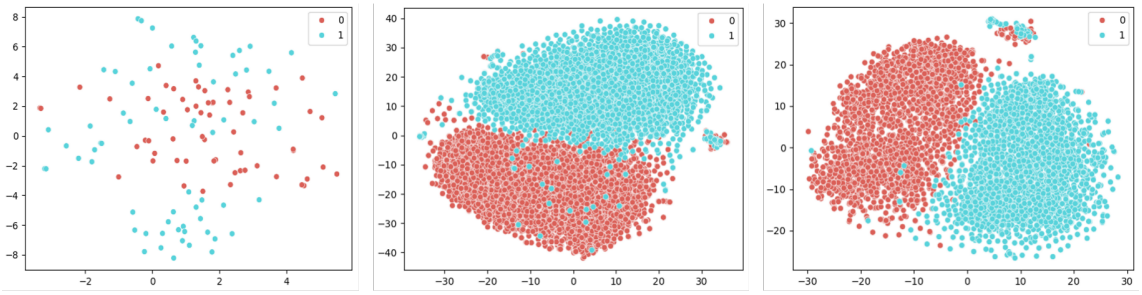


Figure 3.17: t-SNE plot of Breast Cytology Pap Stain dataset(Sakia et al.) Dataset(a) Original Train Set, (b) Augmented Train Set (before selection) (c) Augmented Train Set (after selection; ensemble by three base classifiers)

Table 3.3: A Comparative study of the image quality generated with the proposed Generative model and the state-of-the-art generative models on JUCYT dataset

GENERATIVE MODEL	IS	KID	
	Class Average	Benign	Malignant
SynCGAN [22]	2.02 ± 0.05	0.0668 ± 0.0133	0.0631 ± 0.0106
DCGAN [25]	1.93 ± 0.201	0.0786 ± 0.0110	0.0810 ± 0.0115
Proposed	**2.09 ± 0.11**	**0.0280 ± 0.0061**	**0.0294 ± 0.0056**

Table 3.4: A Comparative study of the image quality generated with the proposed Generative model and the state-of-the-art generative models on Saikia et al dataset

GENERATIVE MODEL	IS	KID	
	Class Average	Benign	Malignant
SynCGAN	1.503 ± 0.0197	0.1449 ± 0.0254	0.1217 ± 0.0199
DCGAN	1.48 ± 0.1503	0.1603 ± 0.0247	0.1183 ± 0.0193
Proposed	**1.508 ±0.0537**	**0.1124 ± 0.0218**	**0.0839 ± 0.0168**

$$IS = exp(\mathbb{E}_{x \sim p_g} D_{KL}(p(y|x) \parallel p(y))) \tag{3.13}$$

where, x be a data sample from the distribution and y be the class label. p_g (probability distribution from the generative model) and $D_{KL}(a \parallel b)$ represent the KL divergence between the distributions a and b. $p(y|x)$ represents the conditional class probability distribution, and $p(y) = \int p(y|x)p_g(x)$ defines the marginal class distribution.

FID Score:

$$FID = ||\mu_1 - \mu_2||^2 + Tr(\sigma_1 + \sigma_2 - 2 * sqrt(\sigma_1 * \sigma_2)) \tag{3.14}$$

where, μ_1, μ_2 refer to the feature-wise mean of the real and generated samples respectively. σ_1, σ_2 refer to covariance matrices of real and synthetic samples respectively. The features are calculated by InceptionNet-v3 model, pretrained on ImageNet 1000 class dataset.

Result Analysis

In the proposed Br-BrCyto-GAN model, the first objective is to develop a generative model, which will synthesize high-quality cytology images. A two-step selective model using fuzzy-based ensemble rule and entropy score has been developed for selecting the best synthetic cytology images, which are more similar to the original ones. Also, we have measured the quality of synthetic cytology images, generated by the proposed model.

In this problem, we have used only Inception score(IS) and Kernel Inception Distance(KID) for the quality evaluation of synthetic samples. For these evaluations, the features are extracted by the InceptionNet-v3 model which is pre-trained on the ImageNet dataset.

In Table 3.3 and Table 3.4, we have compared the proposed BrCyto-GAN generative model with other state-of-the-art generative models with respect to image quality matrics. In the proposed generative model, we have got average IS scores 2.09 $\pm$0.11 and 1.508 $\pm$ 0.053, which are 0.16 and 0.03 higher than the DCGAN(Deep Convolutional GAN) [25] model for JUCYT dataset and breast cytology pap Stain dataset respectively.. Also, we have an average KID score of 0.0287 for JUCYT dataset, which is 0.0511 smaller than DCGAN architecture. On the other hand, SynCGAN model [22] generates images with resolution of 256$\times$256 pixels with an average IS score, KID score of 2.02, and 0.06495 respectively. The proposed BrCyto-GAN model which generates super-resolution images of 512$\times$512 pixels and produces better IS and KID scores than the SynCGAN model. Thus from Table 3.3, it is proved that, if we generate images with high resolution, we can get better quality images.

In Figure 3.16 and Figure 3.17, the effect of fuzzy entropy based selection on the overall class-wise data distribution in the training samples is depicted through t-SNE plots for JUCYT and Breast cytology pap Stain dataset respectively. In Figure 3.16(a), it is shown that the class-wise data distribution in the original training samples is not distinguishable, but in Figure 3.16(c) it is shown that the same is

more distinguishable after the selection of the synthetic samples of JUCYT dataset. The features are calculated by DenseNet-169 model for this t-SNE plot.

Previously generated synthetic cytology samples from SynCGAN and BrCyto-GAN are used as an augmented training set for the next-level classification task. Furthermore, the impact of selectively synthesized high-resolution samples from BrCyto-GAN on malignancy identification is analyzed, revealing a greater effectiveness compared to the SynCGAN model.

3.3.5 Synthetic Cytology Image Generation by Domain to Domain Transfer

Unpaired image-to-image translation models convert images from a source domain to a target domain. In breast malignancy identification, Fine Needle Aspiration Cytology (FNAC) is a cost-effective and minimally invasive diagnostic technique. However, the availability of public FNAC datasets is extremely limited. At the same time, automating cytology image analysis requires a large volume of annotated data.

To bridge this gap, our primary objective is to generate synthetic cytology images by translating publicly available breast histopathology samples, which are comparatively more abundant. In this study, we explore traditional image-to-image translation models, such as CycleGAN and Neural Style Transfer, to assess the feasibility of generating synthetic cytology samples from a different domain.

Histopathology to Cytology image translation using CycleGAN:

In the GAN-based image-to-image translation model, we need to pair source and target images, i.e. for the training process target images are needed for corresponding source images. But for the medical imaging domain, specially for the digital pathology domain, target ground truth mask preparation is quite costly. CycleGAN model can work with an unpaired image-to-image translation model to mitigate this issue. Let A be the source domain(Histopathology) and B be the target domain(Cytology). The main target is to learn a mapping $G : A \rightarrow B$, where the distribution of data from $G(A)$ will be similar like the distribution of data from B. In cycleGAN model, there also another mapping F is defined where

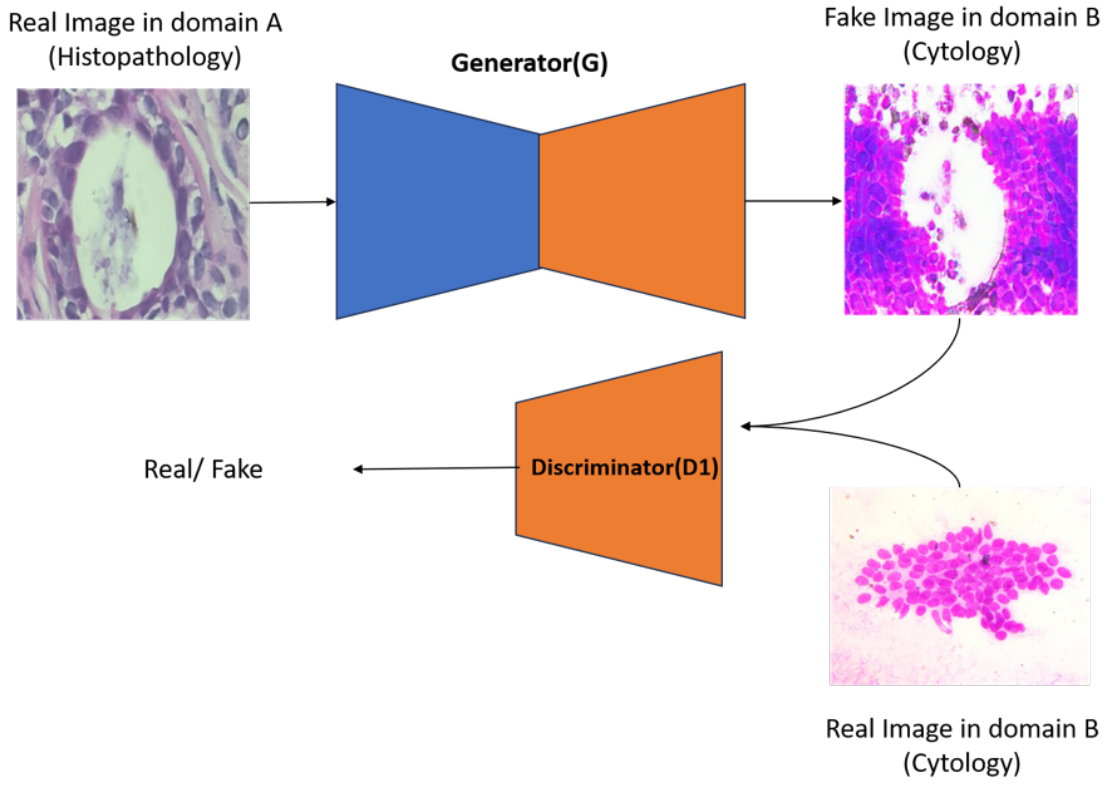


Figure 3.18: Block Diagram of Histopathology to Cytology image translation using CycleGAN

$F : B \rightarrow A$. This inverse mapping is used for generated the synthesis images of source domain. For this study, at testing time, we only translate the images from domain A to domain B. Here, two discriminator and two generator models are used for training. Let D_1 be a discriminative model which will discriminate between $G(a)$ and b where, $a \in A$ and $b \in B$. For the inverse task, let's define another discriminative model D_2, where it's aim to distinguish between a and $F(b)$. Then the adversarial loss for the mapping G is

$$Loss_G(G, D_1, A, b) = Min_G Max_{d_1} \mathbb{E}_b[\log D_1(b)] + \mathbb{E}_a[\log(1 - D_1(G(a)))].$$

The generator G tries to generate fake samples of target domain $G(a)$(i.e. cytology), which will look similar to data from domain B, but D_1 tries to distinguish between the translated sample $G(a)$ and real b.

Another loss, i.e. Cycle Consistency Loss is used on this cycleGAN model, where $Loss_cyc = \mathbb{E}_a[|F(G(a))-a|]+\mathbb{E}_b[|G(F(b))-b|]$. The block diagram of domain(Breast Histopathology) to domain(Breast Cytology) transfer is mentioned in Figure 3.18.

Histopathology to Cytology image translation using Neural Style Transfer:

Neural style transfer [47] is a type of image translation technique, where the synthetic samples are generated by combining the content of source images and the style of target images. In this study, we have used histopathology images as content images and tried to capture the style of cytology images. So, as for the style images, we have randomly chosen the images from the cytology samples.

The features of the content and style images are extracted using a pre-trained VGG-19 model. The loss is computed as a weighted sum of the style loss and content loss. i.e. $Loss_{total} = \alpha Loss_{content} + \beta Loss_{style}$ where $Loss_{content} = \frac{1}{2}\sum(F - P)^2$ where P and F are the feature of the target and content images respectively. The $Loss_{style}$ measures the difference in style between the generated and style image using Gram matrices [30].

Experimental Setup: Here we have used two datasets from two different domains: (a) JUCYT and (b) BreakHis(400X magnification). For experimental study, we have used 169 samples, among which 75 and 94 are Benign and Malignant samples respectively. Also we have used 1820 breast histopathology samples from BreakHis dataset for this study. The details of both datasets are mentioned on Chapter 2.

At training time, the hyper-parameters like batch size, number of epochs, learning rate are set as 1, 100, 0.0002 respectively. PatchGAN discriminator and ADAM are used as discriminator model and optimizer respectively for training cycleGAN model. The model is trained on NVIDIA GTX Geforce 970 GB GPU with Intel core-i5 and RAM 16 GB. The model is trained separately for benign and malignant classes. The training time required for each class is approximately 6 hours.

Result Analysis of Domain-to-Domain Transfer:

In this study, we have explored image-to-image translation model, to generate synthetic breast cytology images, from the publicly available breast histopathology image datasets. In cytology, images have many cellular objects like nuclei, cytoplasm, red blood cells, etc present. and they are located in proper biological order. Also, for the histopathology images, there are some tissue regions present, like stromal region, mitosis, tumor regions, etc. So, from the perspective of morphological information, they are totally different. In this work, we are trying to

translate from one domain to another domain by generative models. On Figure 3.19 and Figure 3.20, we have shown synthetic Benign and Malignant cytology samples respectively, which are generated by CycleGAN models. Also, it is shown from which histopathology image the synthetic images are transformed. We have made a comparative study concerning the quality between real data and synthetic data distributions. Here we have used two traditional quality matrices FID(Frechet Inception Distance)[121] score and KID(Kernel Inception Distance)[12] score values, to check how much synthetic samples are similar to the original cytology samples. In Table 3.5, we have calculated the FID and KID scores between real cytology and fake cytology samples, generated by CycleGAN model. Also in Table 3.6, we have calculated these scores between real histopathology and synthetic cytology samples. It was found that the FID and KID scores between real cytology and synthetic cytology samples are lower than the scores between real histopathology and synthetic cytology. It is observed that, though the synthetic samples are translated from breast histopathology samples, the feature similarity is much better with real cytology samples than the histopathology samples. But visually, it is observed that the cytology samples capture the semantic structures of objects of histopathology samples. The synthetic samples by neural style transfer model are displayed in Figure 3.21 and Figure 3.22. Also, we have measured the quality of fake cytology samples generated by neural style transfer with the original cytology images in Table 3.7. It is observed that the synthetic images generated by CycleGAN are better in quality than the samples generated by neural style transfer. By visual interpretation, we can observe that, in the neural style transfer model, the synthetic samples only capture the style of real cytology images, like background and stain colour, but the main morphological informations like cell size, and cluster of nuclei are not retained, so it is not like an original cytology sample.

Table 3.5: Quality metrics between Real Cytology and Fake Cytology images(generated by CycleGAN model)

Class	FID Score	KID Score
Benign	203.904	0.0163 ± 0.0295
Malignant	143.429	0.0018 ± 0.0154

Table 3.6: Quality metrics between Real Histopathology and Fake Cytology images(generated by CycleGAN model)

Class	FID Score	KID Score
Benign	277.589	0.0968 ± 0.0419
Malignant	222.420	0.0582 ± 0.0423

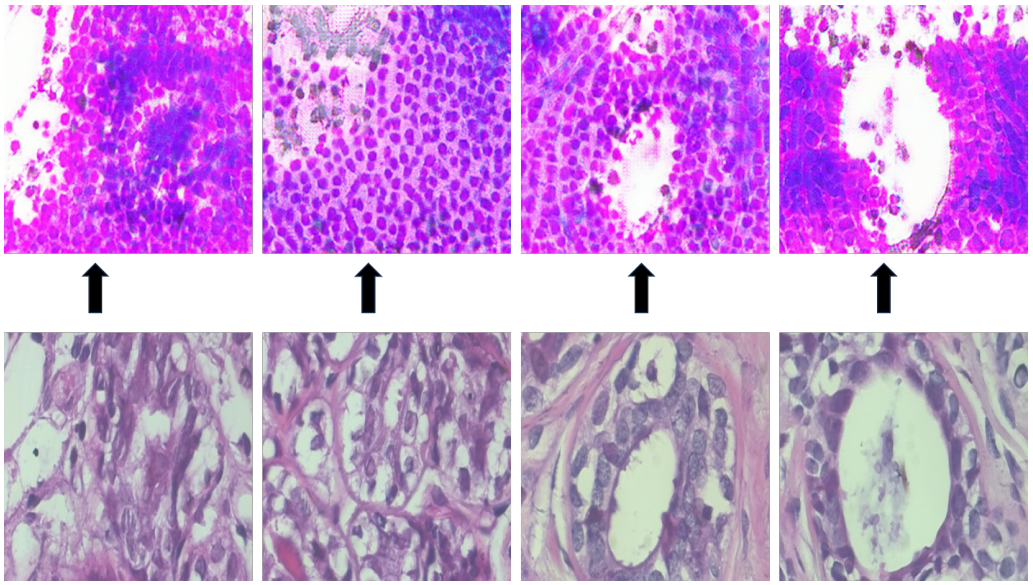


Figure 3.19: Synthetic Benign Cytology images by CycleGAN model. The second row indicates the histopathology images(Source Domain) and the first row indicates the corresponding synthetic cytology images(Target Domain)

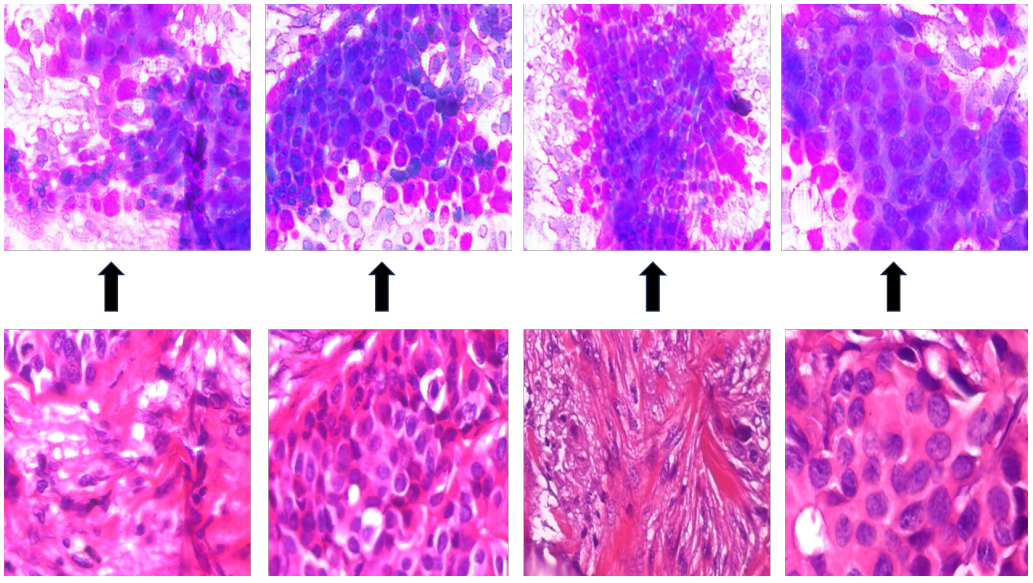


Figure 3.20: Synthetic Malignant Cytology images by CycleGAN model. The second row indicates the histopathology images(Source Domain) and the first row indicates the corresponding synthetic cytology images(Target Domain)

3.4 Discussion

In the context of malignancy identification, the limited availability of training data makes data augmentation essential for addressing class imbalance and improving

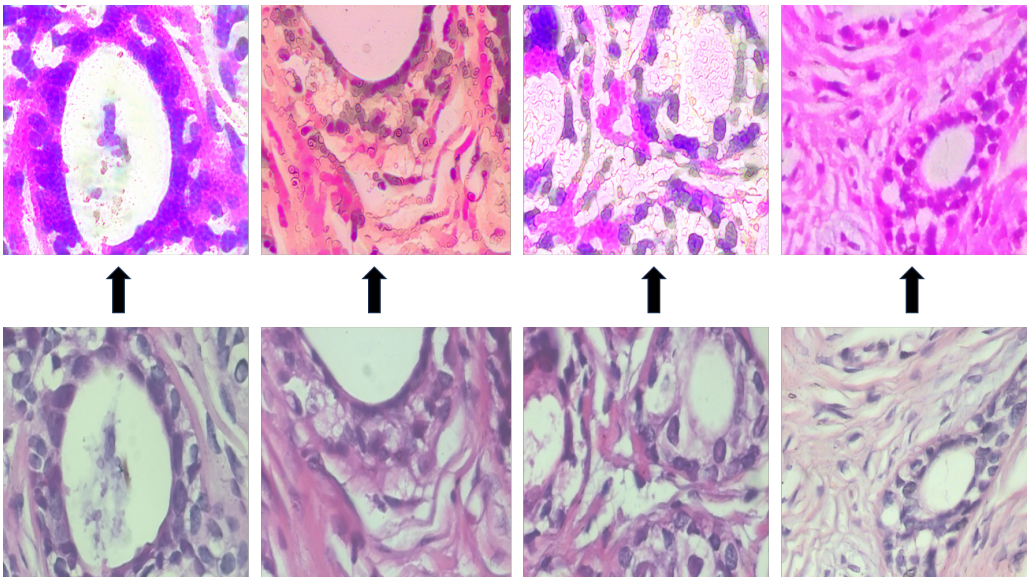


Figure 3.21: Synthetic Benign Cytology images by Neural Style Transfer model. The second row indicates the histopathology images(Content Image) and the first row indicates the corresponding synthetic cytology images

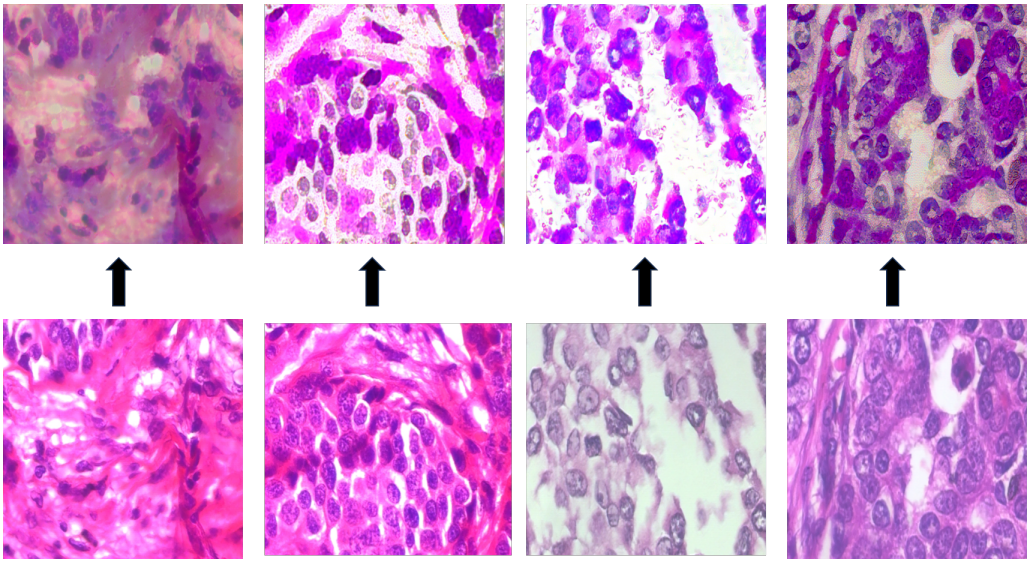


Figure 3.22: Synthetic Malignant Cytology images by Neural Style Transfer model. The second row indicates the histopathology images(Content Image) and the first row indicates the corresponding synthetic cytology images

classification performance. Here, we discuss traditional data augmentation techniques, such as affine transformation and region of interest selection using fuzzy membership rules. Additionally, we explore deep learning-based hybrid models like SynCGAN and Br-CytoGAN for generating synthetic cytology samples. Furthermore, we introduce a selective data augmentation pipeline that employs a fuzzy entropy-based ensemble method to identify and include only those synthetic images in the training set that closely resemble the original ones. Additionally,

Table 3.7: Quality metrics between Real Cytology Samples and Fake Cytology Samples generated by Neural Style Transfer

Class	FID Score	KID Score
Benign	217.58	0.1126 ± 0.0324
Malignant	199.483	0.0763 ± 0.028

we investigate the feasibility of generating cytology samples from other pathology samples, such as histopathology, using the CycleGAN model.

Chapter 4

Breast Cytology Image Segmentation

4.1 Introduction

Image segmentation is a fundamental task in computer vision that divides an image into meaningful, distinct regions or segments, simplifying its representation and enhancing its interpretability for analysis. Its primary goal is to accurately identify and delineate objects or areas of interest within the image. This process is essential for understanding cellular structures and plays a crucial role in diagnosing cancer. The nuclei are segmented to extract nucleus-centric features to classify benign and malignant cells. Breast cytology image segmentation presents several challenges due to the complex nature of the images and inherent variability in the data. The key challenges involved in this segmentation process are as follows: In the breast cytology images, a wide variety of cell shapes, sizes and arrangements are present. Sometimes some overlapping nuclei, and indistinct cell boundaries add complexity for cell segmentation task. In the ductal carcinoma images, dense cluster of overlapping cell in some regions make it difficult to distinguish individual nuclei or cytoplasmic regions. During last few decades, researchers are trying to develop many computer vision techniques for breast cytology image segmentation by solving these challenges. Previously some traditional techniques like clustering, thresolding, etc. are used for segmentation purposes. In the era of deep learning, various CNN models Like UNet [87], SegNet [9], SwinUnet [15], etc. are incorporated for this task. However, grabbing the suitable image segmentation algorithms fit for the purpose is a challenge to the research community.

In this chapter, we explore various methodologies for breast cytology image segmentation using deep learning techniques. We begin with a traditional semantic segmentation model to generate binary segmentation masks. Next, we investigate a fuzzy rank-based ensemble rule for segmenting breast cytology images. The experimental results are validated not only on breast cytology samples but also on digital pathology images from other organs, such as the cervix, etc.. Additionally, we propose a transformer-based framework for breast cytology image segmentation, incorporating a Faster R-CNN-based localization model for cellular object detection. These localized objects are then segmented using the Segmentation Anything Model (SAM). Also, we review previous works in the field of breast cytology image segmentation discussed in this chapter.

4.2 Literature Survey on Breast Cytology Image Segmentation

In the last few decades, various works are reported for breast cytology image segmentation. The researchers have explored traditional unsupervised techniques and also some deep learning approaches for that task. Kowal et al. [56] have implemented deep learning based framework for the maligancy identification from breast cytology samples. Firstly the nuclei are segmented by the combination of marker-controlled watershed algorithm and U-Net architecture, and further some morphological and textural features are extracted from each segmented nuclei. Finally it is classified by traditional machine learning classifiers. The cell nuclei segmentation has achieved F1 score 90% and 86% for benign and malignant cells respectively, on the breast cytology image dataset. Wolberg et al. [116], reported an accuracy of 97.5% for non-overlapped nuclei segmentation on the breast cytology dataset. Usage of compact Hough transform for nuclei segmentation [75] and generalized Hough transform with deformable models [64] are found in literature. Hough transform was also adopted by Hrebien et al. [40] followed by an automatic nuclei localization method based on (1+1) search strategy. To segment nuclei, a combination of active contour model, watershed and grow-cut algorithm was used. But the technique proved to be ineffective for overlapping nuclei. Also, false circles were created that could not be resolved. Marker controlled watershed segmentation was studied by Yang et al. [119]. George et al. [32] suggested

an automated method for nuclei segmentation of breast FNAC images. They extracted Y component of YCbCr color space for grey level conversion followed by Hough transform to detect circular shaped structures. To eliminate false circles, Otsu's thresholding method was applied. To detect nuclei boundaries by avoiding over-segmentation marker controlled watershed transform was used. Seed-based region growing and moving k-means was propounded by Isa et al. [43] to determine the stages of cancer. Filipczuk et al. [28] approached with three level binarization algorithm by extracting the luminance component using the formula L = 0.2126R + 0.7152G + 0.0722B. Initial segmentation was done using adaptive thresholding. Second level involves clustering algorithms such as k-means, FCM, and Gaussian mixture models (GMM) to partition the image into nucleus, cytoplasm and background using different color channels as features. In the final level, they combined two segmented images using an AND operator to give precise definition of the boundaries of the image. But this method suffers from two major limitations. The need for determining optimal parameters and issues associated with unsupervised clustering restricted its use to practical purpose.

4.3 Semantic Segmentation using Deep Learning

4.3.1 Cytology Image Segmentation by Traditional CNN models

In the recent year, CNNs have performed effectively in cytology image segmentation. Semantic segmentation is a trending technique that involves classifying every pixel of an image using machine learning or deep learning models. Various supervised semantic segmentation models like U-Net, SegNet, and PSPNet have been widely utilized for this task. In this study, we have explored these traditional CNN architectures and have generated binary segmentation masks.

UNet Architecture (Ronneberger et al. [87]): UNet is one of the widely used CNN model for biomedical image segmentation. The primary objective of this architecture was to overcome the challenge posed by the limited availability of annotated data in the medical field. It consists of a contracting path and a symmetric expanding path. The encoder layers in the contracting path extract contextual features and decrease the input's spatial dimensions, whereas the decoder layers in

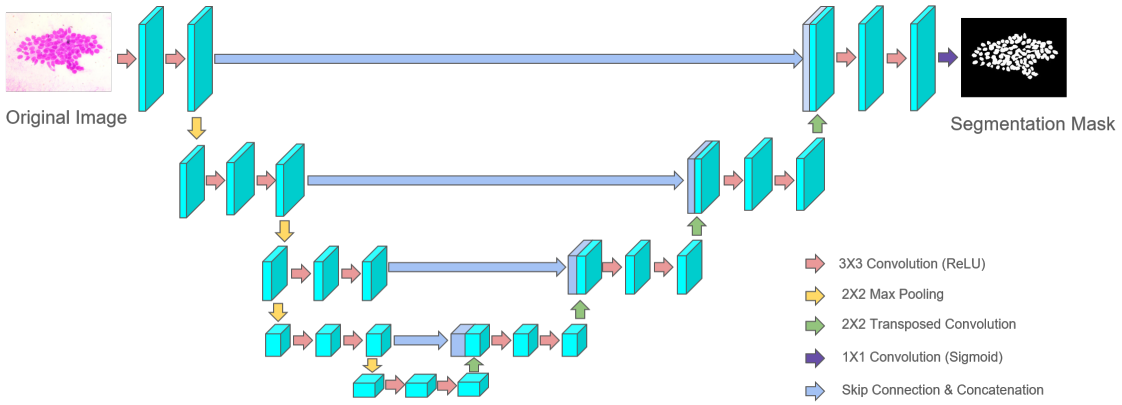


Figure 4.1: Architecture of UNet Semantic Segmentation Model

the expansive path reconstruct the segmentation map by integrating this encoded information with features from the contracting path through skip connections. The encoder block consists of two 3×3 convolution layers, followed by a ReLU activation layer and a 2×2 max pooling layer. The decoder reconstructs the spatial resolution using transposed convolutions (up-convolutions). Each upsampling step is followed by concatenation with the corresponding feature map from the encoder (via skip connections), which helps retain spatial information lost during downsampling. The block diagram of UNet architecture is mentioned in Figure 4.1.

SegNet Architecture (Badrinarayanan et al. [9]): SegNet is a well known semantic segmentation model, which consists of encoder-decoder networks, followed by pixel wise classification layer. The encoder network comprises 13 convolutional layers, aligning with the initial 13 convolutional layers of the VGG16 architecture. Batch normalization is applied after each convolutional layer to enhance training stability and speed. The decoder network comprises 13 layers, each corresponding to a specific encoder layer, facilitating symmetric architecture. The decoding process employs upsampling techniques to progressively reconstruct the feature maps to their original spatial resolution. The block diagram of SegNet model is described in Figure 4.2.

PsPNet(Zhao et al. [126]): Pyramid Scene Parsing Network or PSPNet is a CNN architecture, which has been designed to improve semantic segmentation by effectively capturing both local and global contextual information. The input image is processed using a pretrained ResNet model with a dilated network strategy, enabling detailed feature extraction. This results in a final feature map that is 1/8

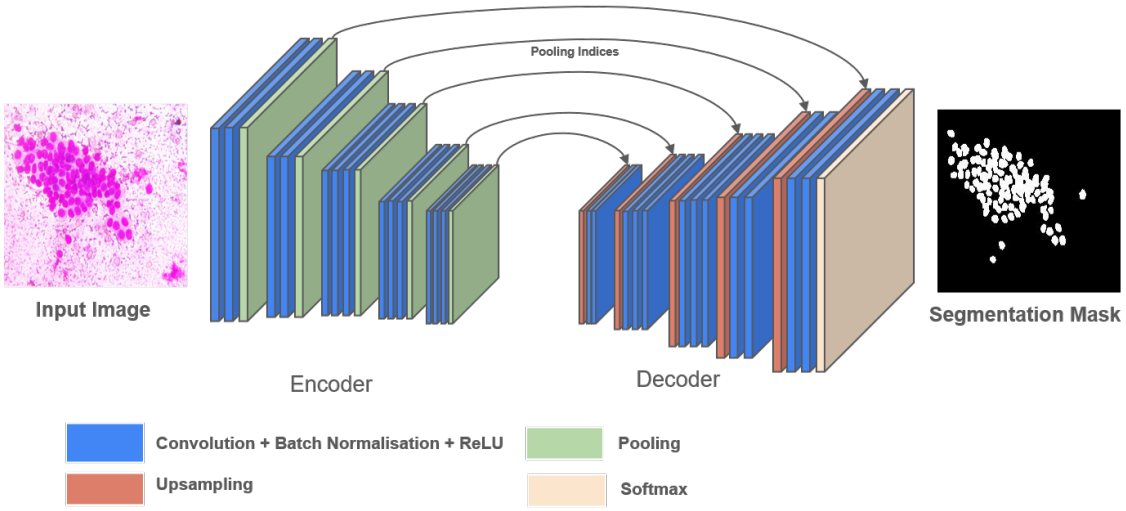


Figure 4.2: Architecture of SegNet Semantic Segmentation Model

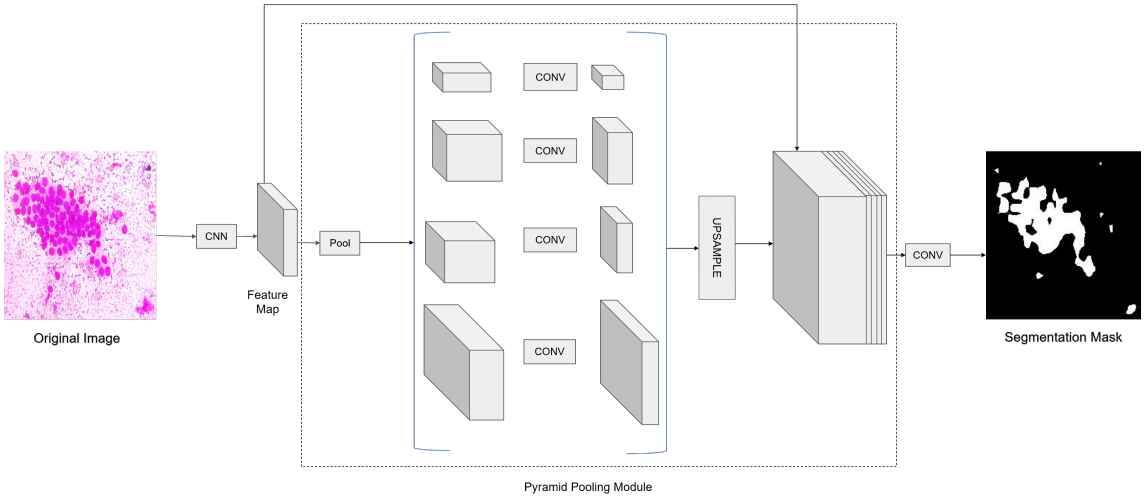


Figure 4.3: Architecture of PSPNet Semantic Segmentation Model

the size of the original image. The Pyramid Pooling Module captures multi-scale contextual information using a 4-level pyramid that pools features from the entire image, half, and smaller regions. These pooled features are fused as global context and concatenated with the original ResNet feature map. The block diagram of PSPNet model is described in Figure 4.3.

Training Process: The cytology images are resized to 224×224 pixels. During CNN (Convolutional Neural Network) training, key hyperparameters are set as follows: batch size = 16, number of epochs = 1000, and learning rate = 0.0001. The ADAM optimizer is used for all networks. The models are trained on NVIDIA-P800 16GB GPU and implemented on PyTorch environment. Training loss is computed

using the negative log-likelihood estimation method. The best-performing models are saved based on the minimum validation loss. By utilizing the best-trained models of SegNet, UNet, and PSPNet, we have evaluated the performance of semantic segmentation at the next level.

At the time of segmentation, each pixel is classified by the semantic segmentation models, ultimately forming a binary segmentation mask for the breast cytology image.

Let $I_{(224\times224)}$ be the cytology image and M_1, M_2, M_3 be the UNet, SegNet, and PSPNet models, respectively. Therefore, for each pixel (x, y) of the image $I_(224 \times 224)$, we have evaluated the class specific scores $\{ y_1^j, y_2^j, ..., y_C^j \}$ with respect to each model M_j for all j and C be the no of classes (for HErlev C=5 and for JUCYT C=2).

$\{ y_1^j, y_2^j, ..., y_C^j \} \leftarrow M_j(I_(224 \times 224(x, y)))$

The pixel (x, y) is predicted as $argmax(\{ y_1^j, y_2^j, ..., y_C^j \})$

4.3.2 Breast Cytology Image Segmentation by Ensemble Rule

The ensemble rule is a machine learning technique where decisions of multiple models are combined using a rule or mathematical model with an objective to achieve an improved decision. In this study, we explored a fuzzy-based ensemble rule for the semantic segmentation of breast cytology images. Additionally, we conducted a comparative analysis with traditional ensemble methods, such as average probability. majority voting, etc..The details of the ensemble rules applied for semantic segmentation are discussed in the sections 4.3.2.1 and 4.3.2.2.

In this work, we employ the previously mentioned CNN models—SegNet, UNet, and PSPNet—as base classifiers within the next-level ensemble framework. The best-trained models from prior evaluations are utilized for this purpose. The class-specific probability distribution for each pixel in the test set is computed using the SOFTMAX function: where, $SOFTMAX(y) = \frac{e^y}{\sum e^y}$,where y represents the class-specific output score from the corresponding CNN model. These probability distribution values serve as inputs for the next-level ensemble task.

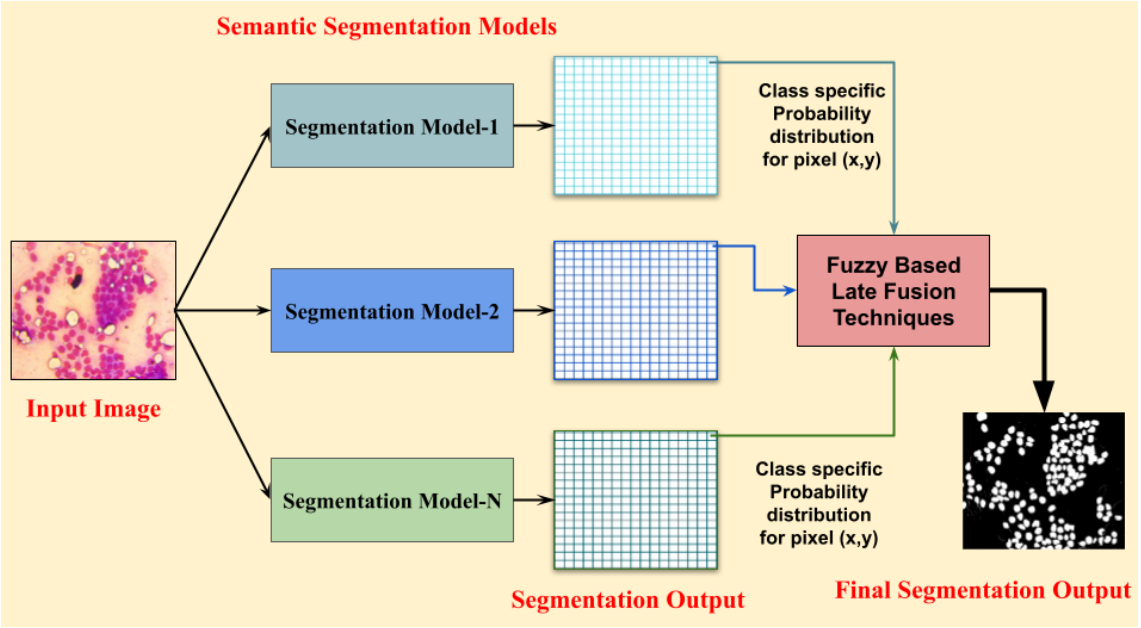


Figure 4.4: Outline of proposed Fuzzy based Late fusion rule for cytology image segmentation

Let M_j be the supervised segmentation model, where $j = 1, 2, 3, ..N$ and N be the total number of base segmentation models. Assume that, $\{ P_1^j, P_2^j, ..., P_C^j \}$ be the set of class specific probability distribution values of the pixel co-ordinate (x, y) for the test image I with respect to segmentation model M_j and C be the total number of class, where $\sum_{i=1}^{C} P_i^j = 1$, $j = 1, 2, .., N$.

4.3.2.1 Fuzzy Rank based Ensemble Technique

Fuzzy rank-based ensemble [21] is a fusion technique where the confidence scores of base learners are fused by membership functions. The two sets of fuzzy rank for segmentation model M_j be $\{ r_1^{j_1}, r_2^{j_1}, ..., r_C^{j_1}\}$ and $\{ r_1^{j_2}, r_2^{j_2}, ..., r_C^{j_2}\}$, which are generated by two non-linear functions such that, $r_k^{j_1} = 1 - tanh(\frac{(P_k^j - 1)^2}{2})$ and $r_k^{j_2} = 1 - exp(\frac{-(P_k^j - 1)^2}{2})$, for each class k. Now assume the set of fused rank scores be $\{ rs_1^j, rs_2^j, ..., rs_C^j\}$ for the model M_j where, $rs_k^j = r_k^{j_1} \times r_k^{j_2}$, for each class k. Let, $\{ fs_1, fs_2, ..., fs_C\}$ be the set of fused scores where $fs_k = \sum_{j=1}^{N} rs_k^j$ for each class k. Finally the predicted class of the pixel (x, y) be **min**$\{ fs_1, fs_2, ..., fs_C\}$.

This operation is executed for all the pixels of the text image $I_{224\times224}$

The flowchart illustrating the fuzzy-based ensemble rule for cytology image segmentation is presented in Figure 4.4.

4.3.2.2 Other Ensemble Techniques

Arithmetic Average rule: Arithmetic average rule [81] is a voting rule, where the mean value of class wise probability distribution values of different models are evaluated.

$P^k_{avg} = \frac{1}{j}\sum_{j=1}^{N}(P^j_k)$, where P^k_{avg} be the average probability of the class k. Therefore the predicted class of the pixel (x, y) be **argmax**$\{ P^1_{avg}, P^2_{avg}, ..., P^C_{avg}\}$.

Geometric Average rule: In geometric average rule, the product value of class wise probability distribution values are evaluated.

$P^k_{prod} = \frac{1}{j}\Pi_{j=1}^{N}(P^j_k)$, where P^k_{prod} be the product of probabilities of the class k. Therefore the predicted class of the pixel (x, y) be **argmax**$\{ P^1_{prod}, P^2_{prod}, ..., P^C_{prod}\}$.

Median value rule: In median value rule [117], the median value of class wise probability distribution values are evaluated.

$P^k_{med} =$ **Median**$\{ P^1_k, P^2_k, P^3_k, ..., P^N_k\}$, where P^k_{med} be the median value of probabilities of the class k. Therefore the predicted class of the pixel (x, y) be **argmax**$\{ P^1_{med}, P^2_{med}, ..., P^C_{med}\}$.

Max Rule In max voting rule [117], the maximum value of class specific probability distribution values are computed. $P^k_{max} =$ **max**$\{ P^1_k, P^2_k, P^3_k, ..., P^N_k\}$, where P^k_{max} be the maximum value of probabilities of the class k. Therefore the predicted class of the pixel (x, y) be **argmax**$\{ P^1_{max}, P^2_{max}, ..., P^C_{max}\}$.

Min Rule In min voting rule [117], the minimum value of class-specific probability distribution values are computed. $P^k_{min} =$ **min**$\{ P^1_k, P^2_k, P^3_k, ..., P^N_k\}$, where P^k_{min} be the minimum value of probabilities of the class k. Therefore the predicted class of the pixel (x, y) be **argmax**$\{ P^1_{min}, P^2_{min}, ..., P^C_{min}\}$.

Borda Count(BC) rule: In BC rule [66], first the class specific probability distribution values of the pixel (x, y) with respect to the model M_j, are arranged in descending order.

$\{ r_1^j, r_2^j, ..., r_C^j\} \leftarrow$ **Descending**$_C\{ P_1^j, P_2^j, ..., P_C^j \}$ where, r_k^j be the rank of the class k with respect to classifier j after arranging the distribution values in descending order. Finally the voting score (v_k) is calculated for the class k by $v_k = \sum_{j=1}^{N}(C - r_k^j)$, for each class k. Therefore the predicted class for the pixel (x, y) be **max**$\{ v_1, v_2, ..., v_C\}$.

4.3.3 Experimental Setup:

Dataset Descriptions: For this study(both individual CNN and ensemble rules), we have used a subset of JUCYT database comprising 62 cytology images, including 39 non-cancerous (benign) and 23 cancerous (malignant) images. The data collection technique (see Figure 2.1) and annotation process (see Figure 2.6) have already been mentioned in the Chapter 2. The manually preparated segmentation masks have consist of two classes: one is foreground and another one is background. For the experiment, it was split into 4:1 ratio, i.e. 50 images for training and 12 images for testing.

Due to the limited availability of publicly accessible breast cytology databases, we applied semantic segmentation models to cytology images from different organs like cervix. Here we have used the Herlev pap smaer dataset for this task (detail is described in Section 2.3). It is divided into training, testing, and validation sets in a 3:1:1 ratio, with 550 images for training, 182 for testing, and 185 for validation. The dataset includes five distinct labels for semantic segmentation. Each image is annotated with five different colors, representing different regions of the Pap smear image: red indicates the background, dark blue represents the cytoplasm, light blue denotes the nuclei, and so on.

4.3.4 Result Analysis

Table 4.1 presents the segmentation performance of these traditional models on two cytology datasets. For the JUCYT dataset, UNet outperforms other two semantic segmentation models. In contrast, for the Herlev dataset, PSPNet performs better than both SegNet and UNet. These CNNs are used as base models for the next-level ensemble task. The primary objective of the proposed Fuzzy based ensemble

model is to perform semantic segmentation by leveraging the combined performance of multiple base semantic segmentation models. On the JUCYT dataset, the combination of UNet and SegNet yielded a mean IoU of 83.79%, reflecting an approximate 5.9% increase in segmentation performance following the proposed ensemble technique. Due to the limited data available for JUCYT, results are reported on the validation set, specifically for the joint performance of UNet and SegNet. We achieved a mean IoU of 84.27% from the combination of UNet and PSPNet on the Herlev dataset, marking a 0.93% improvement over the best-performing base model. Additionally, results indicate that at the base level, PSPNet performs the worst on the JUCYT dataset, failing to capture the structural patterns of breast cytology images effectively. The segmentation masks produced by the base semantic segmentation models are illustrated in Figure 4.5.

We have also conducted a comparative study of traditional ensemble techniques based on their semantic segmentation performance on the cytology dataset. Tables 4.2 and 4.3 present the segmentation performance of various ensemble techniques, achieved through the combination of different base segmentation models. Additionally, the segmentation outputs from different ensemble techniques are illustrated in Figures 4.6 and 4.7.

Table 4.1: Segmentation Performance of Base Semantic Segmentation Model

Dataset	Model	Mean IoU
HErlev	U-Net	81.58
	Seg-Net	71.77
	PSP-Net	83.34
JUCYT	U-Net	77.88
	Seg-Net	66.3
	PSP-Net	56.8

Table 4.2: Different Fusion Rule based performances on HErlev dataset(U= U-Net, S= Seg-Net, P=PSP-Net)

MODEL	Average Probability	Geometric Mean	Median	Maxrule	Minrule	BC-Rule	Fuzzy Rank based Voting
U+S+P	83.42	83.41	82.98	82.62	82.92	82.90	83.38
U+S	75.91	77.77	75.91	72.25	79.52	75.20	75.19
U+P	84.11	84.01	84.11	83.39	84.05	83.06	**84.27**
P+S	82.68	82.56	82.68	82.62	82.19	79.24	82.64

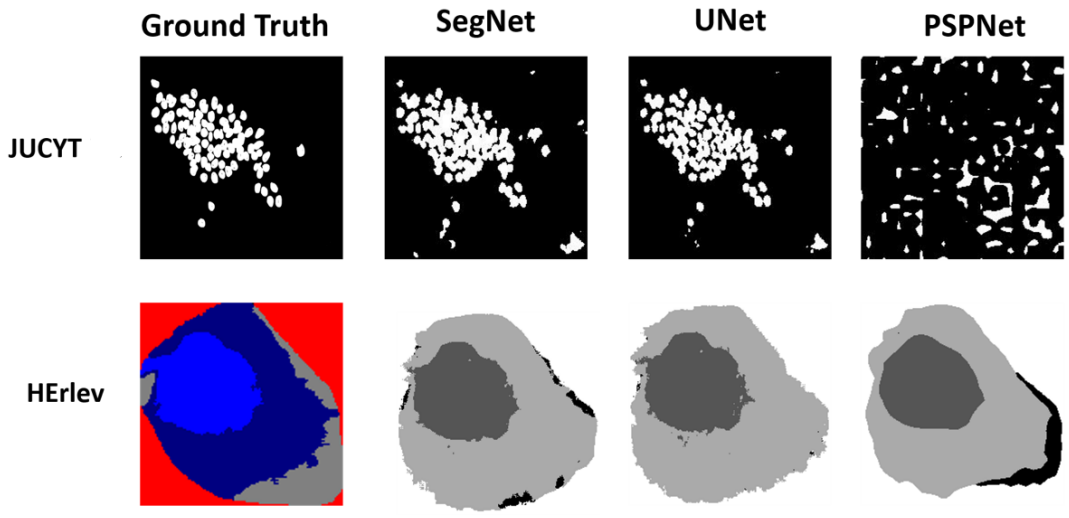


Figure 4.5: Output segmentation masks evaluating by base semantic Segmentation models

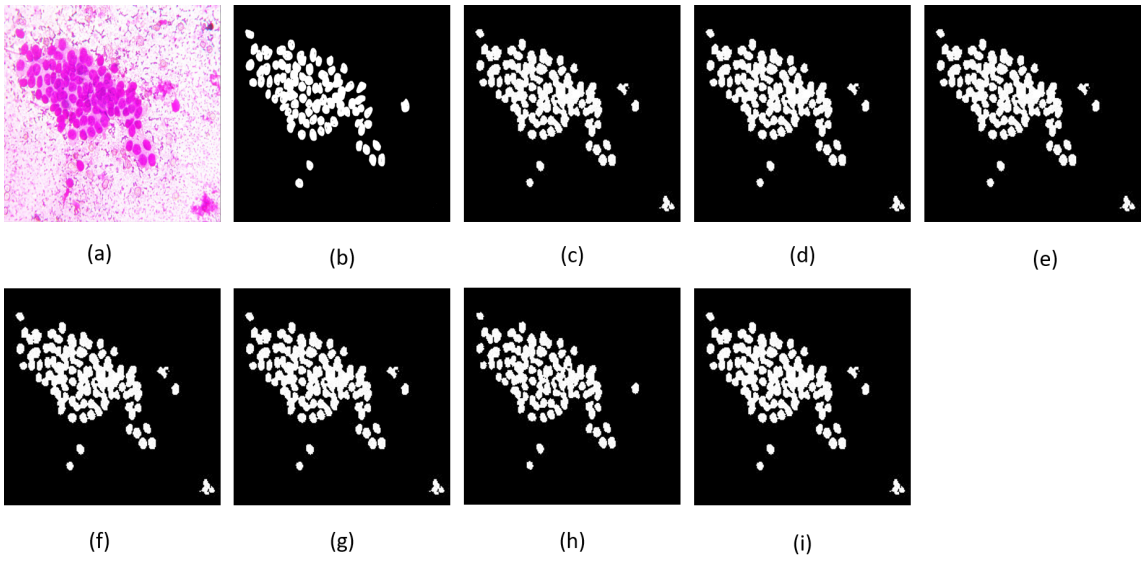


Figure 4.6: Segmentation masks of different fusion techniques(Late fusion of UNet and SegNet) of JUCYT dataset (a) Original Image, (b) Ground Truth, (c) Arithmetic Average, (d) Geometric Average, (e) Median Rule, (f) Max Rule, (g) Min Rule, (h) Borda Count Rule, (i) Fuzzy Rank based Voting rule

4.4 Breast Cytology Image Segmentation using Localization and Segmentation Anything Model

In this study, our primary objective is to localize foreground objects, such as nuclei and cytoplasm, in cytology images using a deep learning based localization model. The next objective was to segment the localized objects using the Segment Anything Model (SAM). Here, we used Faster R-CNN for object localization and

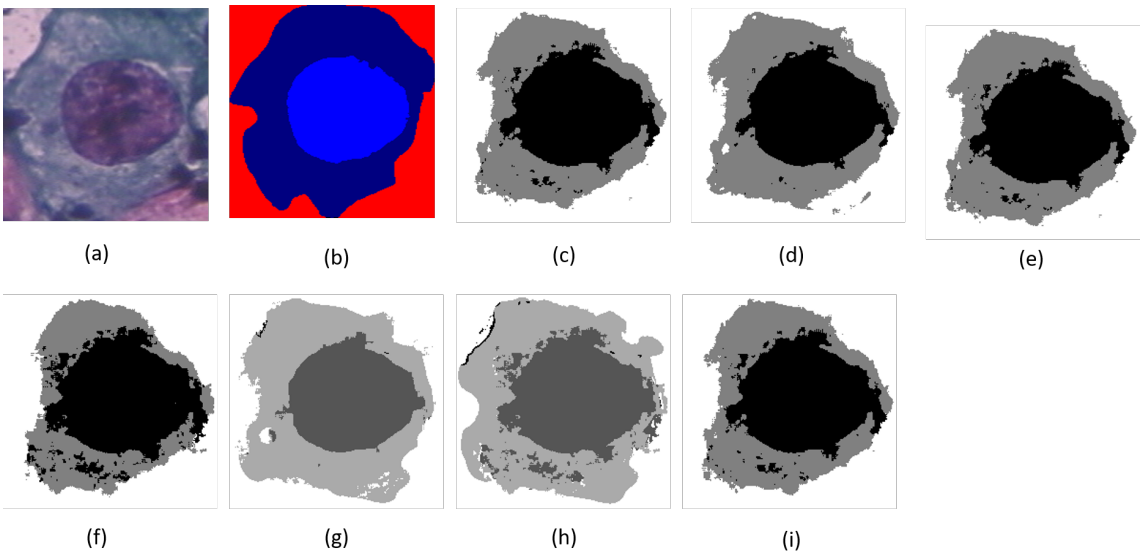


Figure 4.7: Segmentation masks of different fusion techniques(Late fusion of UNet and SegNet) of HErlev Dataset(a) Original Image, (b) Ground Truth, (c) Arithmetic Average, (d) Geometric Average, (e) Median Rule, (f) Max Rule, (g) Min Rule, (h) Borda Count Rule, (i) Fuzzy Rank based Voting rule

Table 4.3: Different Fusion Rule based performances on JUCYT dataset(U= U-Net, S= Seg-Net), P= PSP-Net

MODEL	Average Probability	Geometric Mean	Median	Maxrule	Minrule	BC-Rule	Fuzzy Rank based Voting
U+S+P	70.84	70.24	70.82	68.19	68.19	70.82	71.25
U+S	74.25	74.25	74.25	79.25	74.25	66.56	**83.79**
U+P	66.42	66.42	66.42	66.42	66.42	56.8	66.42
P+S	61.69	61.69	61.69	61.69	61.69	53.69	61.69

MedSAM to generate segmentation masks from the localized regions. The flow diagram of the proposed work is mentioned in Figure 4.8.

Faster-RCNN for Localization Task: Faster RCNN(Faster Region-based Convolutional Neural Network)(Ren et al. [85]) is a well known CNN model for object detection task. The main component in the Faster RCNN network is the base network which significantly determines the accuracy. In the original Faster RCNN model, ZF(Zeiler and Fergus model) and VGG-16 model are used as the base model. This base model is replaced by more advance architecture Feature Pyramid Network, which increases the accuracy to a great extent. Feature Pyramid Network(FPN) is a convolutional neural network module that takes a single-scale image as input and creates a set of feature maps at multiple scales. FPN consists of mainly three modules that are as follows:

Bottom Up Pathway: Bottom Up Pathway uses a feed-forward network that consists of convolution layers grouped into several blocks at each level. Feature maps

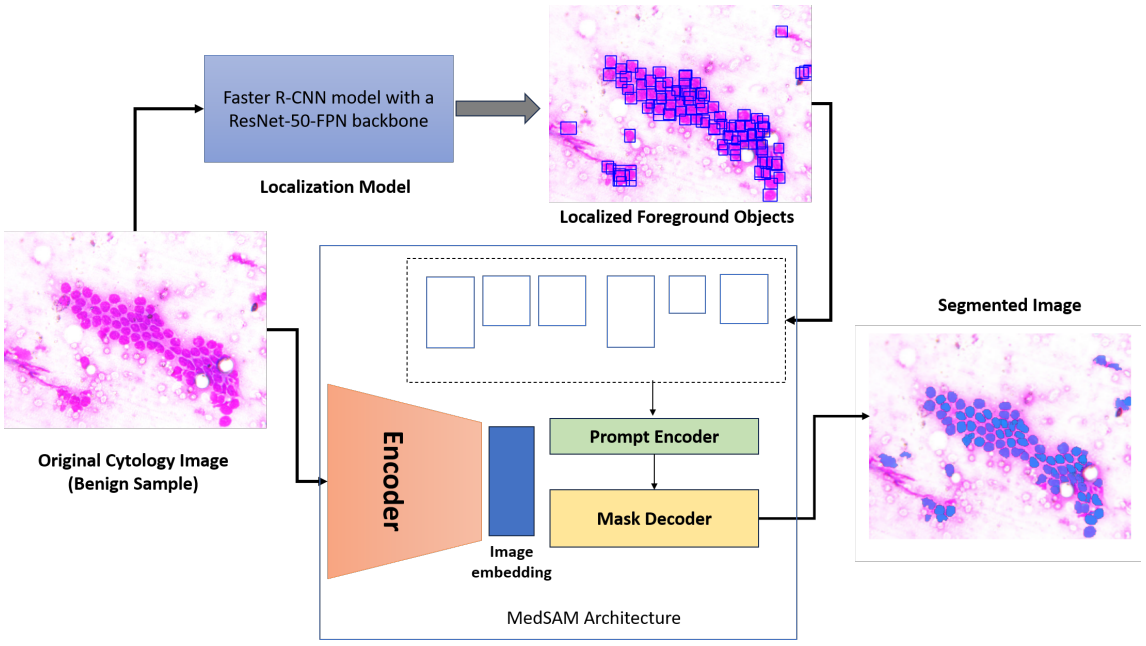


Figure 4.8: Cytology Image Segmentation using MedSAM Model

are obtained from the last layer of each level. The deeper we go into the network; these blocks produce feature maps of lower dimensions than the preceding blocks but have more semantic information.

Top-down Pathway: In top-down pathway, higher resolution feature maps are obtained by up-sampling, from the feature map obtained from the previous level.

Lateral connections: Lateral connections merge feature maps of the same spatial dimensions obtained from the Top-Down and the Bottom-Up pathway.

The Feature Pyramid Network uses ResNet 50 as backbone in the proposed localization method.

MedSAM Model for Segmentation Task : MedSAM (Medical Segment Anything Model)(Ma et al. [67]) is a vision foundation model designed for medical image segmentation task. It is inspired by Meta's Segment Anything Model (SAM) but adapted for the unique challenges posed by medical imaging, such as CT, MRI, and histopathology images. Also it has followed the same architecture of SAM model, consist of image encode, prompt encoder and mask decoder. The primary task of image encoder is to convert high dimensional image embedding from the input image. As a backbone model, pre-trained ViT(Vision Transformer) base is used as image encoder. The ViT base consist of twelve transformer layers, comprising multi

headed self attention block with each block. The prompt encoder transforms the bounding boxes which are evaluated by Faster-RCNN model by the previous stage, into feature representations via positional encoding. It is a lightweight architecture which consist of two transformer layers. Finally the mask decorder combines the image embeddings from the encoder with the promts from prompt encoder. The outputs of the decoder is give the output of the segmentation mask.

Experimental Setup: Here, we used the JUCYT dataset that was previously used for the segmentation task with traditional semantic segmentation models (see Section 4.3) with the same data distribution. The manual annotation for the supervised localization task is made by the experts. Details process of manual bounding box prepartion is mentioned in the Chapter 2(see Figure 2.5).

Training Process:

During the training of Faster-RCNN, the images from the training set and their corresponding bounding box labels for cellular objects are input into the model. The images are resized to a dimension of 256×256. The losses for each iteration are computed for both the Region Proposal Network (RPN) and the ROI Head Detector. The RPN returns a combined loss from both the classification and regression components. For the regression, smooth L1 loss is applied, while cross-entropy loss with logits is used for classification. Stochastic Gradient Descent (SGD) is employed as the optimizer, with a momentum of 0.9. Hyperparameters, including the learning rate, number of epochs, and batch size, are set to 0.001, 50, and 4, respectively. The F1 score for the validation set is computed after each epoch, and the model with the highest F1 score is selected as the best localization model.

The MedSAM model is trained on the RGB image samples and the corresponding masks are resized into 256×256 pixels and the hypermeters like optimizer, batch size, learning rate, weight decay, number of epoch are set to Adam, 2, 0.00001 and 0.01, 500 respectively. Here, two types of loss is calculated one is binary cross-entropy loss, and another is dice loss. The final loss is calculated by the unweighted sum of these two losses.

4.4.1 Result Analysis

Figure 4.9 presents examples of breast cytology samples, illustrating the process from automated localization to segmentation. Here we have made some comparative studies between proposed and state of art localization and segmentation models. Table 4.4 shows that the F1 score of Faster RCNN-ResNet50 FPN is 84% on the JUCYT dataset, outperforming other state-of-the-art localization models. In the proposed Faster-RCNN and MedSAM based automated semantic segmentation model, we have achieved accuracy 84.23%, which is much better than traditional semantic segmentation models like UNet, SegNet and also transformer based semantic segmentation model like SwinUnet(see Table 4.5).

Table 4.4: Comparative study of state of art Localization Models

Model	Mean IoU	Precision	Recall	F1 Score
Yolov3 [127]	0.45	0.56	0.52	0.54
Faster RCNN MobileNet FPN [76]	0.58	0.67	0.63	0.64
Faster RCNN ResNet50 FPN [110]	0.70	0.87	0.82	**0.84**

Table 4.5: Comparative study among state of art semantic segmentation model

Model	Mean IoU
SegNet	66.3
UNet	77.88
SwinUNet	76.82
Proposed (Faster RCNN + MedSAM)	84.23

4.5 Discussion

In this chapter, we review various studies on breast cytology image segmentation. We have employed traditional semantic segmentation models to prepare segmentation masks, and among all the models, UNet outperforms others on the breast cytology dataset. Additionally, we found that a fuzzy-based ensemble of UNet and SegNet models performs better than other ensemble methods for semantic segmentation. We also implemented a transformer-based semantic segmentation model, where cellular objects are initially located using a traditional Faster-RCNN model. The predicted bounding boxes are then used as prompt encoders for the

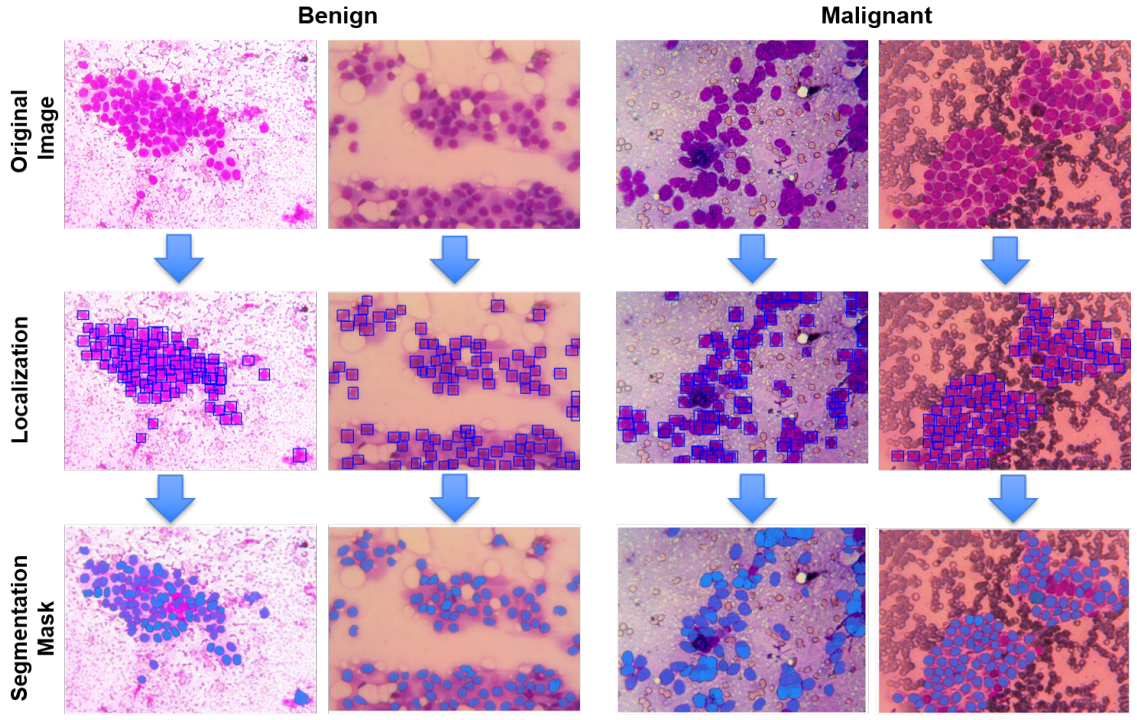


Figure 4.9: Some Examples of Automated Localization to Segmentation Mask Preparation

subsequent MedSAM model. This approach provides more accurate segmentation than other models. However, a key drawback of this model is that the segmented masks are generated at a lower resolution, and overlapping cells are not adequately addressed.

Chapter 5

Breast Cytology Image Classification and Development of Web Application

5.1 Introduction

Ductal carcinoma of the breast is a deadly disease, and it's early diagnosis can significantly reduce the mortality rate. Over the past few decades, researchers have been striving to develop automated systems for breast cancer diagnosis. Street et al. [99] developed a commercially available breast cancer diagnosis system, Xcyt, which utilizes machine learning techniques. In the current era of deep convolutional neural networks (CNNs), we have aimed to develop a low-cost, automated computer-aided diagnosis (CAD) system using deep learning techniques. Our approach explores traditional CNN models along with ensemble techniques involving multiple CNNs to enhance the performance of the image diagnosis system. Additionally, we have employed optimization technique for selecting important features and some data augmentation techniques to further improve diagnostic accuracy.

In this chapter, we discuss traditional CNN architectures for malignancy identification of breast. We introduce the Gaussian Copula rule as an ensemble model and the artificial electric field algorithm as an optimization technique for diagnostic purposes. Additionally, we analyze the impact of data augmentation, demonstrating its effectiveness in improving automated diagnosis. Also, we have evaluated the classification performance of several traditional machine learning models

trained on features extracted from the binary segmentation masks of breast cytology samples. Furthermore, we have developed a publicly available web application for the automated diagnosis of breast cytology image samples. The details of our approach are described in the following sections.

5.2 Literature Survey

Deep learning-based techniques like CNNs are successfully implemented to detect breast cancers. Kalita et al. [48] proposed a deep convolutional neural network architecture with 13 layers for diagnosing breast FNAC samples, achieving an 85% accuracy on the test set. Additionally, they developed a mobile application for a Breast Cancer Detection System based on the TensorFlow framework. Żejmo et al. [124] used AlexNet and GoogLeNet by selecting small patches of 256×256 from the large sized images of 200,000 × 100,000 pixels. They reported accuracies of 80% and 83% on two networks, respectively. It was noticed that the accuracy observed in CNN was still lagging behind traditional feature based models. The accuracy was improved by increasing the number of training samples as suggested by Khan et al. [52]. They proposed a transfer learning based classification technique using VGG net, GoogLeNet, and ResNet. For data augmentation translation, color processing, scaling, horizontal and/or vertical flipping, rotation and noise perturbation techniques were used. Features related to circularity, compactness and roundness were extracted using CNN architectures. The classification accuracy was obtained as 97.525%. Garud et al. [31] proposed a classification technique of breast FNAC images by using GoogLeNet architecture. The model is trained on augmented dataset of FNAC images showing a recognition accuracy 89.7% by using 8 fold cross validation technique. Zerouaoui et al.[125] develops and evaluates twenty-eight hybrid architectures combining seven recent traditional CNN architectures feature extractor, and four machine learning models as classifier for the diagnosis of malignancy from breast FNAC dataset. The hybrid architectures using MLP classifier and DenseNet-201 as feature extractor has performed better than the other hybrid models and it has achieved 99% accuracy on FNAC dataset.

5.3 Automated Classification of Ductal Carcinoma of Breast

Over the past few decades, various computer vision techniques have been proposed for the automated classification of ductal carcinoma using breast cytology images. In cytology samples, the primary informative components are cellular objects such as nuclei and cytoplasm. So, firstly the handcrafted features like area, perimeter, eccentricity, compactness, etc. are extracted from the connected components(cellular objects) of the cytology samples. Based on this feature set, classical machine learning classifiers are employed to categorize the samples as benign or malignant. As classifier, some traditional machine learning models like decision tree, random forest, K-NN(K-Nearest Neighbour), SVM(Support Vector Machine) are used. The classifiers are learned based on the extracted features, and their corresponding parameters are fixed during this training process. In recent years, deep learning techniques have emerged as powerful alternatives that do not rely on handcrafted feature extraction. Deep neural networks, particularly Convolutional Neural Networks (CNNs), automatically learn relevant features directly from the raw input images during the training process. Popular CNN architectures include different versions of ResNet [39], DenseNet [41], InceptionNet [100], etc. The convolutional layers in these models apply filters (or kernels) to the input images, with the initial layers capturing low-level features such as edges and textures, while deeper layers learn high-level representations like shapes and complex patterns. Pooling layers, such as Max Pooling, are used to reduce the spatial dimensions of the feature maps, helping the model to generalize better. After passing through a series of convolutional and pooling layers, the extracted features are flattened into a one-dimensional vector representing high-level abstract features. This vector is then fed into one or more fully connected layers for final classification. During training, CNNs use backpropagation to adjust the model weights, and optimizers (e.g., Adam or SGD) are employed to minimize a predefined loss function. Several hyperparameters—such as the number of epochs, learning rate, and batch size—are set before training, and the best-performing model is typically selected based on the minimum validation loss. To enhance classification performance, several techniques such as ensemble learning, optimization algorithms, and data augmentation are commonly employed. In ensemble learning, well-known approaches include bagging, boosting, average probability, and majority

voting, which combine multiple models to achieve more robust and accurate predictions. Optimization techniques involve fine-tuning deep learning models using specific algorithms to improve convergence and performance. These methods help in extracting more discriminative features, thereby improving classification accuracy. Additionally, to further boost classifier performance, synthetic samples can be incorporated into the training set through data augmentation strategies. This enhances the model's ability to generalize and can lead to improved classification results, especially when dealing with limited or imbalanced datasets.

Some metrics are used for evaluting the classifier performances:

Confussion matrix: It provides a detailed breakdown of the model's predictions compared to the actual ground truth labels. For the binary classification, there are four terms like True Positive(TP), False Positive(FP), True Negative(TN), False Negative(FN).

Precision: It messured the model's reliability when it predicts the positive class. $Precision = \frac{TP}{TP+FP}$

Recall: It messures the model's ability to correctly identify all actual positive cases. $|Recall = \frac{TP}{TP+FN}$

F1 Score: It's evaluated by the harmonic mean of Precision and Recall. $F1 = 2 * \frac{Precision*Recall}{Precision+Recall}$

Accuracy: It is measured by the true prediction over the total number of samples. $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$

Matthews Correlation Coefficient (MCC): $MC = \frac{(TP*TN)-(FP*FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$

5.4 Different Methodologies for Breast Cytology Image Classification

5.4.1 Breast cytology image classification from segmented images:

In this thesis, the main objective was to develop automated system for the diagnosis of ductal carcinoma from breast cytology images using deep learning techniques. Previously many traditional machine learning based techniques [6, 55] have been reported for its diagnosis. Here the models are learned with respect to the cell morphology of cytology samples, which are considers as features.

In the proposed model, the binary segmented masks are prepared using fuzzy rank-based ensemble approach, which combines UNet and SegNet models(details are mentioned in Section 4.3.2.1). Subsequently, handcrafted features [55] of cellular objects—such as area, perimeter, eccentricity, convex area, and compactness—are extracted. To ensure dimensional consistency across samples, only the

top 20 largest objects (based on area) are selected from all connected components.

The details of features are describe in the following:

Area: It is represented by the actual number of pixels in each cellular object.

Perimeter: It is defined by the distance between each adjacent pair of pixels along the border of the cellular object.

Eccentricity: "The cellular regions are predominantly elliptical in shape, and their eccentricity is defined as the ratio of the distance between the foci to the length of the major axis.

Convex Area: It is defined as the number of pixels in the convex image, which represents the convex hull of the object.

Compactness [18]: It describes the proximity of a cell's boundary to its center.

The block diagram of proposed classification technique through features extraction is mentioned in Figure 5.1.

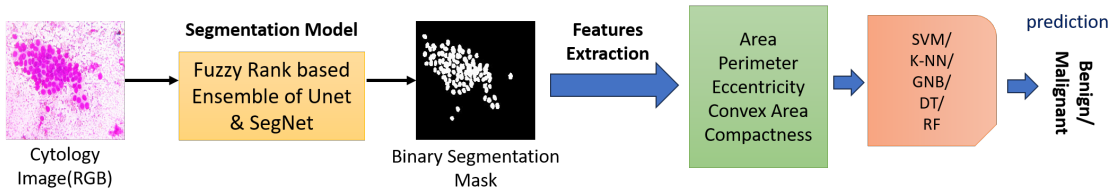


Figure 5.1: Block Diagram of Breast Cytology Image Classification through Features Extraction

Experimental Setup:

For this experiments we have used a subset of JUCYT dataset, which consist of 105 breast cytology samples, of which 60 and 45 are benign and malignant, respectively. The 50 samples are used for training and the rest of the samples are used for testing the machine learning models. Here we have used traditional machine learning models like Support Vector Machine(SVM), Decission Tree(DT), Gaussian Naive Bayes(GNB), K-Nearest Neighbour(K-NN), Random Forest(RF) as classifier. The experiments are conducted on the system with 8Gb RAM and i3 processor.

Result Analysis

The experimental results for the JUCYT dataset using machine learning models are presented in Table 5.1. From this table, it can be concluded that the SVM classifier achieves the highest accuracy of 72%, outperforming the other machine learning models. In contrast, the Random Forest classifier performs the worst among the tested models. In addition to accuracy, other performance metrics such as precision, recall, and F1 score are also calculated.

Table 5.1: Classification Performances on JUCYT dataset By Machine Learning based Classifiers

Model	Accuracy	Precision	Recall	F1-Score
SVM	**72**	74.5	69.09	66
DT	63.33	66	69.09	63.5
K-NN (K=3)	64	66	64	63
RF	56	57	57	56
GNB	69.09	72	70	63.5

5.4.2 Breast Cytology Image Classification by Copula based Ensemble technique: GC-EnC

In this work, we have explored the potential of Copula-based ensemble of CNNs over individual classifiers for malignancy identification in breast cytology images.

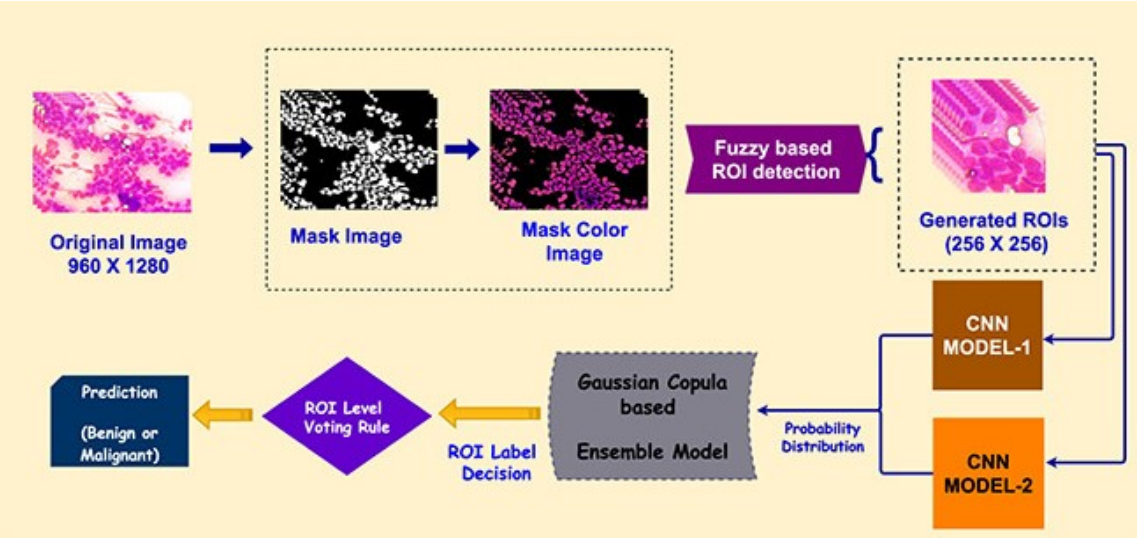


Figure 5.2: Block Diagram of Proposed GC-Enc Model

The Copula-based model(GC-EnC) that integrates three best-performing CNN architectures from different versions of ResNet, DenseNet and InceptionNet, is proposed. Also, the limitation of small dataset is circumvented using a Fuzzy template based data augmentation technique(see Section 3.3.2) that intelligently selects multiple region of interests (ROIs) from an image. The selection process of the best-performing CNNs is described in the following paragraph. Finally the images are classified by ROI voting rule. The block diagram of proposed GC-EnC model described on Figure 5.2.

Experimental Setup:

The experiments are primarily conducted on JUCYT(Breast Cytology) dataset. For Gc-Enc model, we have used 156 breast cytology images(consisting 77 benign and 79 malignant images) from the JUCYT dataset (details provided in Chapter 2), captured from 30 FNAC slides. To validate the ensemble[92] performance further, three additional datasets like BreakHis(Breast Histopathology), BI (Breast Histopathology) and SipakMeD (Cervical Cytology) are used. The descriptions of this datasets are briefly mentioned in Chapter 2(see Section 2.3). We introduce the Fuzzy membership guided data augmentation technique to extract ROIs (Figure 3.7) to produce variance of data to train the CNNs. We have used NVIDIA GTX Geforce 970 GPU with Intel core-i5 and 16 GB RAM for training the CNN models. Due to non-availity of standard data distribution, we have splited the dataset on two ways:

1. The datasets are proportionately divided into 3:1:1 for training, test and validation sets respectively.

2. Five fold cross validation is done.

Choice of best CNN models

Three popular CNN models; DenseNet, ResNet and InceptionNet with different architectures are considered to obtain the best trained model and to compute class specific probability distribution values for making a decision in fusion process. Here we have chosen these CNN architectures because they are popularly used for classification of digital pathology images like histopathology [14] and cytology [90] images.

ROIs or sub-images are generated from a whole slide image by using the ***Fuzzy()*** described in Algorithm 1(see Chapter 3). A maximum number of 20 ROIs have been observed for an image. Three sets of ROIs (train, test, and validation set) are generated in this process and are combined into a single set to represent the total number of sub-images of the dataset. These ROIs are labelled with their corresponding original image labelled and also arranged to train, test validation set based on its original images' set.

A set of ROIs, thus generated from a single image, has been labelled according to the label of the original image and is added to the original dataset.

Thus number of samples in train, validation and test set are increased.

The increased number of training samples are then used for training CNN architectures with hyper parameter tuning by augmented validation set. CNN architectures of InceptionNet (V1, V2, V3), ResNet(18, 34, 50, 101, 152) and DenseNet (121, 161, 169, 201) are trained with ROIs of train set.
The models are trained with a learning rate of 0.001, batch size 8, maximum epoch of 200 with an ADAM optimizer(a popularly used optimization algorithm for training CNN models and it is handling sparse gradients on noisy problems if any) and a negative log-likelihood estimation based training loss. During training input images are resized [33] to 224$\times$224 for ResNet and DenseNet and 299$\times$299 for InceptionNet to make it compatible with their original input sizes [36]. The traditional pre-trained models on ImageNet [82] are employed by changing the

number of classes. Best trained models are then resolved empirically using parameter tuning on validation set of ROIs.

Now, the best trained models of InceptionNet, ResNet and DenseNet are obtained from the ROIs of the train set.

Now the class specific probability distribution values of each ROI of the test set, are calculated by the *SOFTMAX()* function with respect to the best trained models of InceptionNet, ResNet and DenseNet. Where, $SOFTMAX(\mathbf{x})=\frac{e^{\mathbf{x}}}{\sum e^{\mathbf{x}}}$ for input vector **x**.

Now the ROIs are predicted with respect to maximum of probability distribution values. After that the images of original test set are predicted using ***ROIVOTE()*** function (see Algorithm 3). From the Algorithm, we can compute the number of TRUE predictions of the test set. Finally we have calculated the accuracy of every InceptionNet($\mathbb{I}$-V1/V2/V3), DenseNet($\mathbb{D}$-121/161/169/201), ResNet($\mathbb{R}$-18/34/50/101/152). The maximum accuracy obtained on test set among different InceptionNet models is used for the next level of classifier fusion. Similarly the best models of ResNet and DenseNet are chosen by this proposed ROI voting algorithm.

Algorithm 3: Prediction of an image by ROI voting

Input:
V $\leftarrow$ set of predicted classes of ROIs of Image I
$p \leftarrow$ voting percentage
$I_{LC} \leftarrow$ labeled prediction of test image I
Output : TRUE or FALSE
ROIVOTE(**V**, p ,I_{LC})
$M_{count} \leftarrow$0 //Number of Predicted *Malignant* ROI
$I_{OP} \leftarrow$ Prediction output of I
$m \leftarrow |\mathbf{V}|$ // number of elements in **V**
for $j = 0$ to m **do**
    if $V(j)$ ***is*** *Malignant* **then**
        $M_{count} \leftarrow M_{count}$ + 1
    end if
end for
if $(M_{count}/m) \geq 0.01 \times p$ **then**
    $I_{OP} \leftarrow$ *Malignant* **else** $I_{OP} \leftarrow$ *Benign*
end if
if $I_{LC} = I_{OP}$ **then**
    return TRUE **else** return FALSE
end if

Similarly to find the best trained model using 5-fold cross validation, we split the entire dataset into 5 number of folds, and training process is repeated 5 times by taking 4 folds for training and one fold for validation in a round robin fashion. Similarly the dataset of 5 folds are extracted by using ***Fuzzy()*** function. Then the previously obtained best performing models of InceptionNet, ResNet and DenseNet are applied on the 5 fold dataset. After that the performance of the trained model on the validation set for each fold is computed by previously mentioned ROI vote approaches. Finally, the average performance on the validation set is measured. In this 5 fold cross validation technique the validation set is considered as the test set.

A probabilistic Copula is then used to create a classification framework that amalgamates best architectures which is discussed in the subsequent section.

Classifiers fusion using Probabilistic Copula

A brief description of Copula

From the last few years, Copula based models have been introduced in various areas of computer vision such as segmentation [62, 59], regression [77] etc. In the present framework, Gaussian Copula(GC) [78] is employed for ensembling probability scores of best CNN models.

The Copula is defined as a mapping from univariate marginal probability distribution to multivariate distribution. Let, U_1,U_2,...,U_n be n uniform random variables marginally distributed on [0,1] defined on a joint probability distribution function (pdf), such that,

$$\mathbf{C}(u_1, u_2, .., u_n) = \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2, .., U_n \leq u_n) \tag{5.1}$$

where, **C** is the Copula function defined over **C**:$[0,1]^n \rightarrow [0,1]$, and $\mathbb{P}$ is the probability function. By Sklar's theorem [78], let, $F_1(x_1), F_2(x_2), ..., F_n(x_n)$ be n univariate marginal cumulative probability distributions of $x_1, x_2, ..., x_n$ respectively. $F(x_1, x_2, ..., x_n)$ be N-dimensional cumulative probability distribution function over real-valued random variables $x_i \in [-\infty, +\infty]$,$\forall i \in 1, 2, ..., n$.Thus,

$$F(x_1, x_2, .., x_n) = \mathbf{C}(F_1(x_1), F_2(x_2), ..., F_n(x_n)) \tag{5.2}$$

Using Eq. 5.2, joint pdf can be derived as,

$$\begin{aligned} f(x) &= \frac{\delta^n}{\delta x_1, \delta x_2, ..., \delta x_n}\mathbf{C}(F_1(x_1), F_2(x_2), ..., F_n(x_n)) \\ &= \mathbf{c}(F_1(x_1), F_2(x_2), ..., F_n(x_n)) \prod_{i=1}^{n} f_i(x_i) \end{aligned} \tag{5.3}$$

where, **c** is the Copula density function such that

$$\mathbf{c}(u_1, u_2, .., u_n) = \frac{\delta^n}{\delta u_1, \delta u_2, ..., \delta u_n}\mathbf{C}(u_1, u_2, .., u_n) \tag{5.4}$$

Suppose, there are t classifiers, and their probability scores are p_i, where $i = 1, 2, ..., t$. From Eq. 5.3, we can derive the joint probability distribution of $p_1, p_2, ..., p_t$ under the class L_c for each c. (L_1= *Malignant* class and L_2= *Benign* class in our case) as,

$$\begin{aligned} &f(p_1, p_2, ..., p_t|L_c) = \\ &\qquad \mathbf{c}(F(p_1|L_c), F(p_2|L_c), ..., F(p_t|L_c)) \prod_{l} f(p_l|L_c) \end{aligned}$$

The fused probability score p_0 is

$$\begin{aligned} p_0 &= g(p_1, p_2, ...p_t) \\ &= P(L_c|p_1, p_2, ..p_t) \propto f(p_1, p_2, ., p_t|L_c)P(L_c) \end{aligned} \tag{5.5}$$

where, $f(p_1, p_2, ..., p_t|L_c)$ is the class-specific joint likelihood probability for the class L_c and $g : [0, 1]^t \rightarrow [0, 1]$ be the mapping for computing fused probability scores.

A GC function can be defined as

$$\mathbf{C}_{Gauss}(u_1, u_2, .., u_n) = \phi_\Sigma(\phi^{-1}(u_1), \phi^{-1}(u_2), ..\phi^{-1}(u_n)) \tag{5.6}$$

and the GC density function is written as,

$$\begin{aligned} \mathbf{c}_{Gauss} &= \frac{1}{\sqrt{det\Sigma}} * \\ exp(-\frac{1}{2} &\begin{pmatrix} \phi^{-1}(u_1) \\ \phi^{-1}(u_2) \\ . \\ . \\ \phi^{-1}(u_n) \end{pmatrix}^T .(\Sigma^{-1} - \mathbf{I}). \begin{pmatrix} \phi^{-1}(u_1) \\ \phi^{-1}(u_2) \\ . \\ . \\ \phi^{-1}(u_n) \end{pmatrix}) \end{aligned} \tag{5.7}$$

where, **I** is the identity matrix, ϕ_Σ is the joint cumulative distribution function(cdf) of a multivariate normal distribution with correlation matrix Σ and ϕ^{-1} is the inverse cdf.

Marginal Density Estimation: When data is fitted to GC, the probability distribution values of data are not used directly [78]. Then cumulative distribution of data are estimated by Kernel Density Estimation(KDE)[122] method. KDE is a non-parametric Kernel Smoothing Marginal Density Estimation [122] function $f(x)$ defined as,

$$f(x) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x - X_i}{h}) \tag{5.8}$$

where, K(x) is the kernel function and $X_1, X_2, ..., X_n$ are n univariate data and $h > 0$ is the smoothing bandwidth.

The cumulative distribution function $F(x)$ of any real values of x is defined as,

$$F(x) = \int_{-\infty}^{x} f(t)dt \qquad = \frac{1}{n}\sum_{i=1}^{n} G(\frac{x - X_i}{h}) \tag{5.9}$$

where, $G(x) = \int_{-\infty}^{x} K(t)dt$

In this step, previously evaluated class specific probability distribution values would be used by each classification model.

The best performing models of DenseNet, ResNet and InceptionNet are selected for next level ensemble by using Gaussian Copula. First, the class specific probability distribution values of train and test set are computed by using ***SOFTMAX*** function. After that, the probability distribution values are passed to ***CopulaFusion*** function for the prediction.
Let $\mathbf{V_k^{(d)}}$ be the set of probability distribution values of class k (B=Benign, M=Malignant) of set d, where, d= train(tr) or test (te) and N be the number of classifiers. Thus, the predicted class of ROIs are calculated as:
$P_{ROI} \leftarrow$ ***CopulaFusion***$(\mathbf{V_B^{(tr)}}, \mathbf{V_M^{(tr)}}, \mathbf{V_B^{(te)}}, \mathbf{V_M^{(te)}}, |N|)$.
From these, the predicted class of an image of the test set will be calculated by the ***ROIVOTE()*** function. Finally, the accuracy of test set is calculated and it is repeated for all the 5-folds.

Results Analysis:

The most appropriate CNNs architectures are identified for the copula ensemble technique (See Algorithm 4) based on the maximum accuracies among different ResNet, DenseNet and InceptionNet architectures using ROI voting scheme as shown in Figure 5.3. The ROC curves (True Positive Rate versus False Positive

Algorithm 4: Copula Based ensemble of CNNs

Input: $\mathbf{V_B}^{(tr)}, \mathbf{V_M}^{(tr)}, \mathbf{V_B}^{(te)}, \mathbf{V_M}^{(te)}, |N|$
Output: P_{ROI}
CopulaFusion($\mathbf{V_B}^{(tr)}$,$\mathbf{V_M}^{(tr)}$, $\mathbf{V_B}^{(te)}$, $\mathbf{V_M}^{(te)}$, $|N|$)
$\rho_{|N|\times|N|} \leftarrow \mathbf{C}_{Gauss}([F(\mathbf{V_B}^{(tr)})]_{|\mathbf{V_B}^{(tr)}|\times|N|})$
$f_{\mathbf{B}} \leftarrow \mathbf{c}_{Gauss}([F(\mathbf{V_B}^{(te)})]_{|\mathbf{V_B}^{(te)}|\times|N|}), \rho) * (\prod_{|N|} \mathbf{V_B}^{(te)}) * prior$ // for binary classification *prior* is chosen 0.5
$\rho'_{|N|\times|N|} \leftarrow \mathbf{C}_{Gauss}([F(\mathbf{V_M}^{(tr)})]_{|\mathbf{V_M}^{(tr)}|\times|N|}))$
$f_{\mathbf{M}} \leftarrow \mathbf{c}_{Gauss}([F(\mathbf{V_M}^{(te)})]_{|\mathbf{V_M}^{(te)}|\times|N|}), \rho') * (\prod_{|N|} \mathbf{V_M}^{(te)}) * prior$
Compute for each $f_{\mathbf{B}} \leftarrow \frac{f_{\mathbf{B}}}{\sum f_{\mathbf{B}}}$ and $f_{\mathbf{M}} \leftarrow \frac{f_{\mathbf{M}}}{\sum f_{\mathbf{M}}}$
if $f_B > f_M$ **then**
  $P_{ROI} \leftarrow$ *Benign* **else** $P_{ROI} \leftarrow$ *Malignant*;
end if
return P_{ROI}

Rate) of the different architectures of ResNet, DenseNet and InceptionNet are also given for all the four datasets (See Figure 5.4 and Figure 5.5). Thus, best models for ensemble of CNNs are $\mathbb{D}$-161, $\mathbb{R}$-101, $\mathbb{I}$-$V3$ for JUCYT, $\mathbb{D}$-161, $\mathbb{R}$-34, $\mathbb{I}$-$V3$ for BreakHis, $\mathbb{I}$-$V3$, $\mathbb{R}$-34, $\mathbb{D}$-161 for SipakMeD, $\mathbb{I}$-$V3$, $\mathbb{R}$-34, $\mathbb{D}$-201 for BI dataset. The findings are validated using both the AUC and Accuracy values (see Figures 5.3, 5.4, 5.5).

Since the proposed recognition performance is based on ROI voting approach, so the selection of appropriate percentage of ROIs is important [31]. For cytology or histopathology dataset, if at least a certain percentages of the total numbers of ROIs are predicted as malignant, then the original image is classified as malignant. This percentages are chosen empirically, for BrekHis, BI, JUCYT, SipakMeD dataset, it is is chosen 5%, 45%, 5%, 35% respectively. Using that voting scheme, we have got the best classification performance of 84.37%, 91.67%, 97.3129%, 98.98% on the JUCYT, BI, BrekHis, SipakMeD datasets respectively. Here, $\mathbb{I}-V3$ is selected for every dataset in ensemble task, so the classification performances of $\mathbb{I}-V3$ model on the validation set are evaluated in different voting percentages from 5% to 50% (see Figure 5.6). The voting percentages are selected with respect to best performance of $\mathbb{I}-V3$ model in different dataset. In this proposed method, only for SipakMeD dataset we have used existing ROIs of the datasets for ROI voting scheme, since the fuzzy template based ROIs generated in our scheme are in the dimension of 256×256 pixels which is not suitable for visualization of the nuclei and cytoplasm part properly. Therefore, the features from those fuzzy ROIs are

not captured properly. So the classification performance is quite less than original ROI. But the copula based ensemble is performing better than individual classifiers in fuzzy ROI based approaches, which is the main contribution of the proposed work. The details results of using fuzzy ROI detection on SipakMeD dataset are mentioned in Table 5.2. Here we selected the best models from DenseNet, ResNet and InceptionNet as $\mathbb{D}$-161, $\mathbb{R}$-34 and $\mathbb{I}-V3$. After Gaussian copula based ensemble, it has achieved 92.8205% in the fusion of $\mathbb{R}$-34 and $\mathbb{I}-V3$. For SipakMeD dataset we have trained the CNN models with respect to five classes, but at the time of test phase we merged it into binary classes.

Table 5.2: Performance comparison of individual CNNs and Gaussian Copula based fusion rule based models on SipakMed dataset.(Here ROIs are selected by proposed *Fuzzy template based ROI detection* technique)

Individual CNN performance		**Copula based fusion performance**	
CNN	**Accuracy**	**Fusion Model**	**Accuracy**
$\mathbb{D}$-201	88.2051	$\mathbb{I}$-V3+$\mathbb{D}$-161	91.2821
$\mathbb{D}$-121	83.5897	$\mathbb{I}$-V3+$\mathbb{D}$-161+$\mathbb{R}$-34	91.7949
$\mathbb{D}$-169	90.7692	$\mathbb{I}$-V3+$\mathbb{R}$-34	**92.8205**
$\mathbb{D}$-161	**90.7692**	$\mathbb{D}$-161+$\mathbb{R}$-34	91.2821
$\mathbb{R}$-34	**90.7692**		
$\mathbb{R}$-18	87.6923		
$\mathbb{R}$-50	90.2564		
$\mathbb{R}$-101	89.2308		
$\mathbb{R}$-152	89.7436		
$\mathbb{I}$-V1	89.2308		
$\mathbb{I}$-V3	**90.7692**		
$\mathbb{I}$-V4	87.1795		

Table 5.3: Classification Accuracy (ACC) in % using Traditional(TDA) and proposed ROI-based Data Augmentation(PRDA) technique

JUCYT			**BreakHis**		
CNN	**TDA ACC(in %)**	**PRDA ACC (in %)**	**CNN**	**TDA ACC (in %)**	**PRDA ACC (in %)**
$\mathbb{D}$-161	59.38	68.75	$\mathbb{D}$-161	71.75	86.16
$\mathbb{R}$-101	53.12	71.87	$\mathbb{R}$-34	73.75	80.51
$\mathbb{I}$-$V3$	62.50	71.87	$\mathbb{I}$-$V3$	74.63	87.57

In Table 5.3, the performance of CNNs using proposed ROI based augmentation technique (PRDA) and traditional data augmentation techniques (mentioned in Chapter 3) (TDA) [95] such as random crop, random horizontal flip, random vertical flip, and random rotation upto 30° during training are compared. All the performances are measured on 3:1:1 data distribution among train, validation and test set.

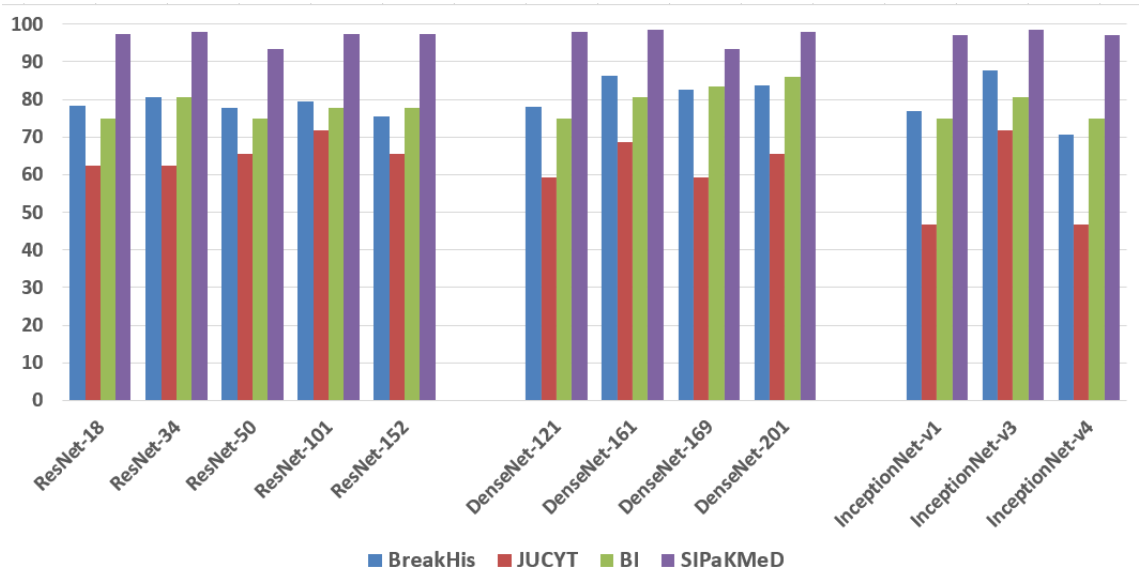


Figure 5.3: Accuracy comparison among different ResNet, DenseNet and InceptionNet architectures on testdata set by ROI voting approaches

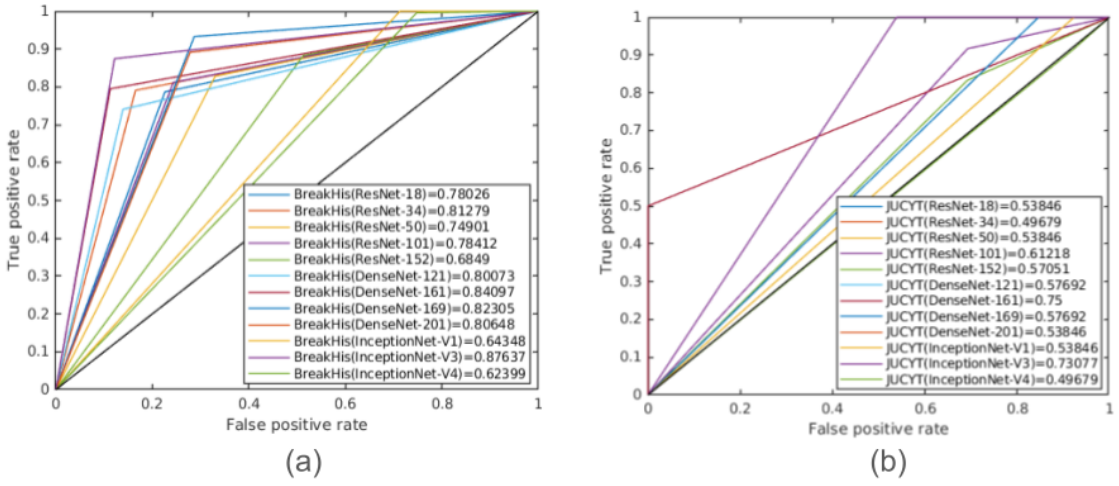


Figure 5.4: ROC Curve with AUC value of (a)BreakHis dataset and (b) JUCYT dataset

It is observed that, ROI-based data augmentation is better suited to capture regional information which boosts the performance of CNNs compared to the traditional one. Maximum accuracies of 87.57% and 71.87% have been observed for BrekHis and JUCYT datasets using $\mathbb{I}$-$V3$ and $\mathbb{I}$-$V3$/$\mathbb{R}$-101 respectively. The results of several combinations of CNN architectures are ensembled using Gaussian Copula function are encapsulated in Table 5.4.

Ensemble of $\mathbb{I}$-$V3$ and $\mathbb{D}$-161 scored the maximum accuracies of 90.96% (BreakHis), 84.37% (JUCYT) and 98.98% (SipakMeD) on test data, where as BI dataset scored

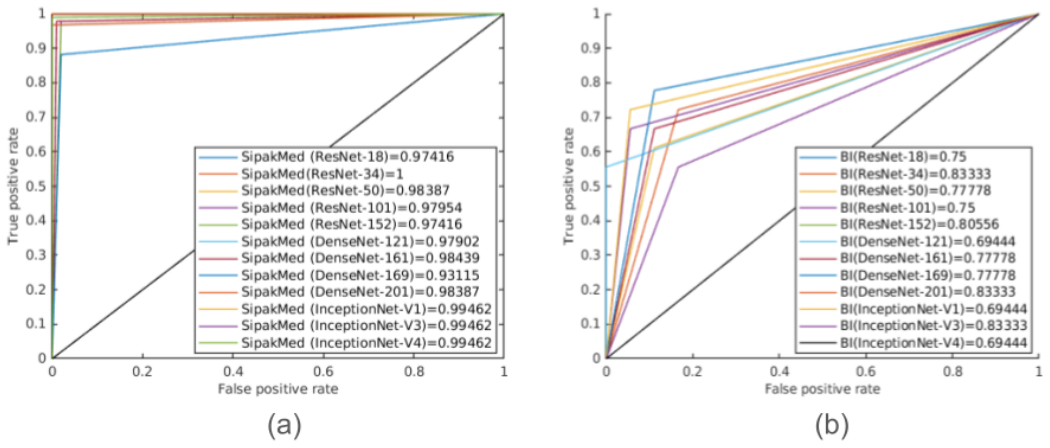


Figure 5.5: ROC Curve with AUC value of (a)SipakMeD dataset and (b) BI dataset

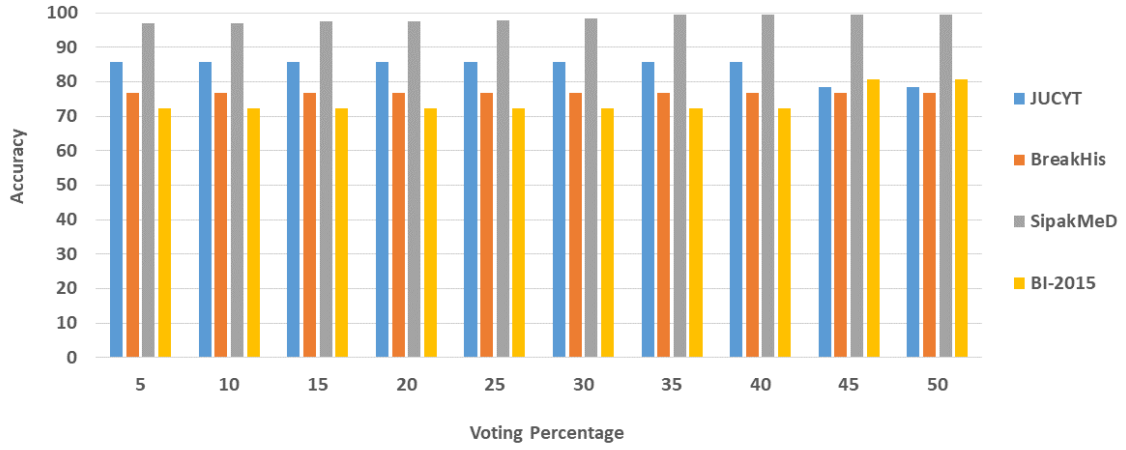


Figure 5.6: Performance of Inception-V3 model in different voting percentage on validation set

91.67% accuracy on the ensemble of $\mathbb{I}$-$V3$, $\mathbb{D}$-201 and $\mathbb{R}$-34. The results unveil the potential of Gaussian Copula to ensemble deep learning based classifiers. The concept is also ratified using 5-fold cross validation as shown in Table 5.5. Maximum average accuracies of 97.32%, 78.31%, 98.46% are observed on BreakHis, JUCYT and SipakMeD datasets respectively on 5-fold cross validation distribution. The other popular ensembling techniques i.e., average probability, majority voting, bagging (bootstrap aggregating) are evaluated and got the maximum average accuracies of 96.78%, 96.81%, 91.03% for BrekHis dataset; 76.99%, 73.37% and 66.67% for JUCYT dataset; 98.26%, 98.36%, 90.31% for SipakMeD dataset; 88.89%, 88.89%, 86.11% for BI dataset(see Table 5.6) which are lower than the

Table 5.4: Performance On Individual CNN and their ensemble on Cytology(JUCYT & SipakMeD) and Histopathology(BREAKHIS & BI) Images.(Accuracy=Acc)

Individual CNN Performance							
JUCYT		BreakHis		SipakMeD		BI	
CNN	Acc	CNN	Acc	CNN	Acc	CNN	Acc
$\mathbb{I}$-V3	71.87	$\mathbb{I}$-V3	87.57	$\mathbb{I}$-V3	98.47	$\mathbb{I}$-V3	80.56
$\mathbb{R}$-101	71.87	$\mathbb{R}$-34	80.51	$\mathbb{R}$-34	97.96	$\mathbb{R}$-34	80.56
$\mathbb{D}$-161	68.75	$\mathbb{D}$-161	86.16	$\mathbb{D}$-161	98.47	$\mathbb{D}$-201	86.11
Ensemble CNN Performance							
JUCYT		BreakHis		SipakMeD		BI	
Fused CNN	Acc	Fused CNN	Acc	Fused CNN	Acc	Fused CNN	Acc
$\mathbb{I}$-V3 + $\mathbb{R}$-101	71.87	$\mathbb{I}$-V3 + $\mathbb{R}$-34	87.9	$\mathbb{I}$-V3 + $\mathbb{R}$-34	**98.98**	$\mathbb{I}$-V3 + $\mathbb{R}$-34	86.11
$\mathbb{I}$-V3 + $\mathbb{D}$-161	**84.37**	$\mathbb{I}$-V3 + $\mathbb{D}$-161	**90.96**	$\mathbb{I}$-V3 + $\mathbb{D}$-161	**98.98**	$\mathbb{I}$-V3 + $\mathbb{D}$-201	88.89
$\mathbb{D}$-161 + $\mathbb{R}$-101	78.125	$\mathbb{D}$-161 + $\mathbb{R}$-34	87.3	$\mathbb{D}$-161 + $\mathbb{R}$-34	98.47	$\mathbb{D}$-201 + $\mathbb{R}$-34	88.89
$\mathbb{I}$-V3 + $\mathbb{R}$-101 + $\mathbb{D}$-161	81.25	$\mathbb{I}$-V3 + $\mathbb{R}$-34 + $\mathbb{D}$-161	88.7	$\mathbb{I}$-V3 + $\mathbb{R}$-34 + $\mathbb{D}$-161	**98.98**	$\mathbb{I}$-V3 + $\mathbb{R}$-34 + $\mathbb{D}$-201	**91.67**

proposed ROI voting based Gaussian copula based ensemble method. Observations from Figure 5.7 also reveal that area under curve (AUC) values for Gaussian Copula ensemble (0.97 for BreakHis and 0.78 for JUCYT) overrule the AUC values for majority voting (0.95 for BreakHis and 0.74 for JUCYT) and average probability-based ensemble (0.96 and 0.77 for BreakHis and JUCYT), from which superiority of the Copula-based ensemble method can be established. Also the ROC and AUC are described for SipakMeD(Fold wise) and BI(standard distribution) at Figure 5.8 and Figure 5.9 respectively. The comparison of the proposed technique with the state-of-the-art methods for BreakHis, SipakMeD, BI dataset are summarized in Table 5.7.

The performance metrices of the best ensemble models of five-fold cross validation on JUCYT, BreakHis, SipakMeD and BI datasets are mentioned at Table 5.8. The best model depicts an average sensitivity and specificity of 73% and 83% for JUCYT; 95% and 98% for BreakHis dataset, 98% in both cases for SipakMeD dataset using 5 folds which indicate a high correct classification rate of the *Malignant* and *Benign* class. On BI dataset the best ensemble model has achieved 100% specificity, i.e. all normal and benign samples are predicted as non cancerous and also there is no false positivity.

Matthews-Correlation coefficient values of 0.94 (BreakHis) indicates a strong correlation between the actual and the predicted class. For cytology images the value is relatively low i.e. 0.57 which reflects that there is a scope to improve by including of more variations of images on training set. The fold wise (standard distribution only for BI dataset) confusion matrices of the best performing ensemble models are described in the Figure 5.10, Figure 5.11, Figure 5.12, and Figure

5.13. As BreakHis dataset does not have any predefined train, test data distribution which makes it difficult to compare with previous methods on even test-bed. So we made a comparison with respect to 5-fold cross-validation. Since cytology dataset is introduced recently; the performance comparisons are unavailable on a uniform platform.

We ensembled the probability distribution values of different classifiers. In some cases, it is observed that ensemble results of two classifiers have performed better than the ensemble of three classifiers as it depends on the joint probability distribution values[88]. The performance of the ensemble model may be the higher or nearer to best performing CNNs [88]. Same observation is recorded on other probability based fusion techniques like average probability, majority voting. Average execution time of 3.2 second/image is taken using for Gaussian copula based ensembling of DenseNet-161 and InceptionNet-v3, whereas the individual networks like DenseNet-161 takes an average of 2.88 seconds/image. Some examples of cytology and histopathology images along with their classification results using individual classifiers and their ensemble techniques are given in Table 5.9.

For the JUCYT dataset, out of 25 test images, there are six instances in which the base classifiers ResNet and InceptionNet predict false, but DenseNet predicts true. Five, two of the six images have been predicted as true by the ensemble techniques based on Gaussian Copula and Average Probability, respectively. On the BreakHis test dataset, it is found that two base classifiers predicted false for 12 images, whereas one base classifier predicted true. The GC ensemble predicted 7 of these 12 images as true, although the average probability rule predicted just one One image from the test set on the BI dataset is predicted true by DenseNet but as false by ResNet and InceptionNet. This image is also predicted as true by the Gaussian Copula ensemble, but it is predicted false by other ensemble techniques such as majority voting and average probability. From the statistics, it can be concluded that even though two classifiers are predicted wrong at the initial phase, the Gaussian Copula-based ensemble technique predicted correctly most of the time, whereas others failed. This indicates the robustness of Gaussian Copula based ensemble technique over others.

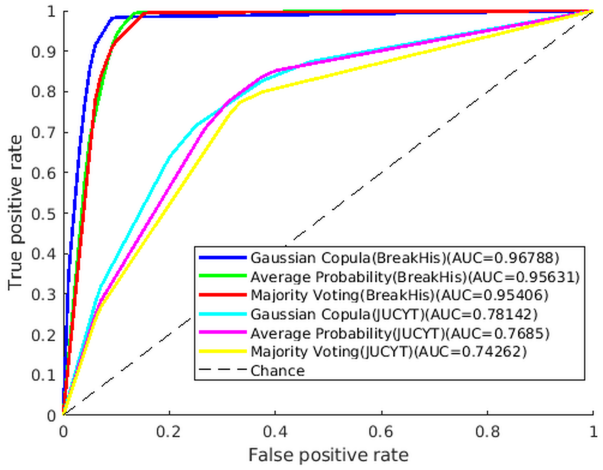


Figure 5.7: Mean ROC curves of different classifier fusion techniques BreakHis and JUCYT for five folds

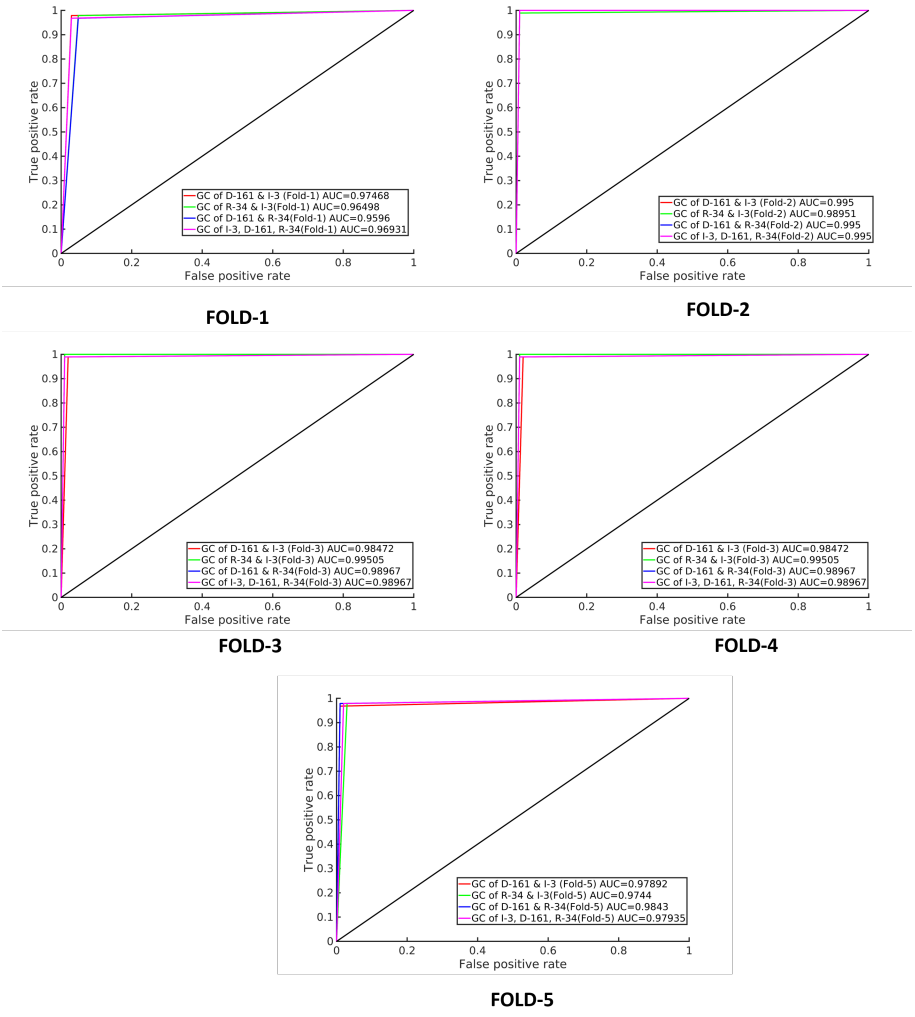


Figure 5.8: ROC curve and AUC values of SipakMeD 5-Fold dataset of all combinations of Copula ensemble

Table 5.5: Comparative study between Gaussian Copula based Ensemble and Traditional Ensemble Methods for JUCYT, BreakHis, SipakMeD dataset (Five fold cross validation data distribution)

Dataset	Classification Rule	Model	Fold-1		Fold-2		Fold-3		Fold-4		Fold-5		Average	
			ROI Voting	Without ROI Voting	ROI Voting	Without ROI Voting	ROI Voting	Without ROI Voting	ROI Voting	Without ROI Voting	ROI Voting	Without ROI Voting	ROI Voting	Without ROI Voting
JUCYT	Individual CNN	𝕀-V3	67.74	61.29	80	73.33	76.67	73.33	71.88	62.5	66.67	72.73	72.6 ± 5.7	68.64 ± 6.2
		ℝ-101	77.42	70.96	60	53.33	70	70	78.13	78	75.76	75	72.27 ± 7.57	69.46 ± 9.6
		𝔻-161	70.97	67.74	80	73.33	76.67	76.67	68.75	65.63	81.82	80	75.65 ± 5.7	72.68 ± 6.1
	Gaussian Copula	𝕀-V3 + ℝ-101	74.19	70.97	76.67	76.67	80	76.67	71.88	68.75	75.76	78.79	75.7 ± 3.0	74.37 ± 4.3
		𝕀-V3 + 𝔻-161	70.97	74.19	86.67	86.67	83.33	80	68.75	59.38	81.82	87.88	**78.31 ± 7.9**	77.63 ± 11.6
		𝔻-161 + ℝ-101	80.65	80.64	76.67	80	76.67	80	65.63	65.63	81.82	84.85	76.29 ± 6.4	**78.23 ± 7.4**
		𝕀-V3 + ℝ-101 + 𝔻-161	74.19	70.97	83.33	80	76.67	76.67	68.75	68.75	78.79	69.7	76.35 ± 5.4	73.22 ± 4.9
	Average Probability	𝕀-V3 + ℝ-101	74.19	70.96	73.33	73.33	73.33	73.33	75	62.5	75.75	78.78	74.32 ± 1.1	71.78 ± 6
		𝕀-V3 + 𝔻-161	70.96	67.74	83.33	80	80	73.33	71.87	59.37	78.78	87.87	**76.99 ± 5.37**	73.67 ± 11
		𝔻-161 + ℝ-101	80.64	80.64	76.67	80	76.67	76.76	65.62	68.75	81.82	75.75	76.29 ± 6.4	**76.38 ± 4.8**
		𝕀-V3 + ℝ-101 + 𝔻-161	64.51	67.74	73.33	67.74	80	76.76	75	68.75	75.75	81.81	73.72 ± 5.71	72.56 ± 6.5
	Majority Voting	𝕀-V3 + ℝ-101 + 𝔻-161	67.74	67.74	73.33	73.33	75	76.67	75	68.75	75.76	78.78	73.37 ± 3.27	73.06 ± 4.9
	Bagging Ensemble	𝕀-V3 + ℝ-101 + 𝔻-161											66.67	
BreakHis	Individual CNN	𝕀-V3	98.05	95.28	96.49	95.54	94.95	92.44	94.1	93.54	97.2	91.92	96.16 ± 1.62	93.75 ± 1.7
		ℝ-34	92.5	86.11	92.48	84.68	93.28	86.84	90.17	87.36	92.72	86.28	92.23 ± 1.2	86.26 ± 1.1
		𝔻-161	98.06	96.94	91.36	83.57	95.52	95.24	95.79	94.66	97.76	94.68	95.7 ± 2.7	93.02 ± 5.4
	Gaussian Copula	𝕀-V3 + ℝ-34	96.94	94.72	97.77	93.59	95.24	95.79	94.38	93.53	97.48	93.59	96.37 ± 1.5	94.25 ± 1
		𝕀-V3 + 𝔻-161	99.44	97.5	96.38	91.92	97.48	96.08	94.66	96.07	98.6	96.64	**97.32 ± 1.8**	**95.65 ± 2.2**
		𝔻-161 + ℝ-34	97.78	94.72	94.99	90.81	95.24	92.44	95.22	93.82	96.92	94.72	96.03 ± 1.2	93.31 ± 1.7
		𝕀-V3 + ℝ-34 + 𝔻-161	99.17	95.56	96.66	94.15	96.36	94.4	95.51	94.94	98.04	94.68	97.15 ± 1.5	94.75 ± 0.6
	Average Probability	𝕀-V3 + ℝ-34	96.67	94.16	95.82	93.59	95.23	92.99	93.82	93.25	96.07	91.59	95.53 ± 1.1	93.12 ± 1
		𝕀-V3 + 𝔻-161	98.05	96.94	95.54	89.97	95.35	96.07	96.62	95.78	98.31	96.63	**96.78 ± 1.4**	**95.08 ± 2.9**
		𝔻-161 + ℝ-34	97.5	94.72	93.87	89.69	94.95	91.87	95.22	93.53	96.07	92.71	95.53 ± 1.4	92.51 ± 1.9
		𝕀-V3 + ℝ-34 + 𝔻-161	97.77	94.72	96.49	95.54	96.35	94.4	95.22	93.53	97.47	94.67	96.66 ± 1.0	94.58 ± 0.8
	Majority Voting	𝕀-V3 + ℝ-34 + 𝔻-161	97.78	95.83	97.77	93.59	96.36	95.79	94.66	93.54	97.48	94.67	**96.81 ± 1.3**	**94.69 ± 1.2**
	Bagging Ensemble	𝕀-V3 + ℝ-34 + 𝔻-161											91.03	-
SipakMeD	Individual CNN	𝕀-V3	96.94	94.89	97.91	95.29	97.94	94.84	97.38	97.38	97.42	92.78	97.52 ± .5	95.04 ± 1.7
		ℝ-34	95.41	94.85	96.86	91.1	97.94	94.33	97.91	94.76	97.42	94.33	97.11 ± 1.1	93.88 ± 1.6
		𝔻-161	96.93	92.86	98.95	91.62	95.88	95.36	98.95	93.72	96.39	94.85	97.42 ± 1.5	93.69 ± 1.6
	Gaussian Copula	𝕀-V3 + ℝ-34	96.43	94.39	98.95	93.19	99.48	95.87	98.43	97.38	97.42	94.84	98.15 ± 1.3	95.14 ± 1.6
		𝕀-V3 + 𝔻-161	97.45	93.88	99.48	96.86	98.45	97.42	98.95	97.38	97.94	94.85	**98.46 ± 0.9**	**96.08 ± 1.7**
		𝔻-161 + ℝ-34	95.92	94.39	99.48	93.19	99.48	96.39	98.95	97.38	97.94	94.84	98.36 ± 1.5	95.24 ± 1.7
		𝕀-V3 + ℝ-34 + 𝔻-161	96.94	94.89	99.48	95.28	98.97	96.91	98.43	97.38	97.94	93.81	98.36 ± 1	95.66 ± 1.5
	Average Probability	𝕀-V3 + ℝ-34	96.43	93.19	98.95	93.19	98.95	95.88	98.43	97.38	97.42	94.84	98.04 ± 1.1	94.9 ± 1.9
		𝕀-V3 + 𝔻-161	97.45	93.88	99.48	95.36	97.94	97.42	98.95	97.38	97.94	94.84	98.36 ± 0.9	**95.78 ± 1.6**
		𝔻-161 + ℝ-34	95.92	93.37	99.48	92.15	98.97	95.36	98.95	97.38	97.94	96.33	**98.26 ± 1.5**	94.92 ± 2.2
		𝕀-V3 + ℝ-34 + 𝔻-161	97.45	94.9	98.95	94.24	97.94	96.39	98.95	96.86	97.94	94.33	98.25 ± 0.7	95.35 ± 1.3
	Majority Voting	𝕀-V3 + ℝ-34 + 𝔻-161	96.43	95.4	99.48	94.76	99.48	95.88	98.43	96.33	97.94	94.33	**98.36 ± 1.3**	**95.34 ± 0.9**
	Bagging Ensemble	𝕀-V3 + ℝ-34 + 𝔻-161											90.31	-

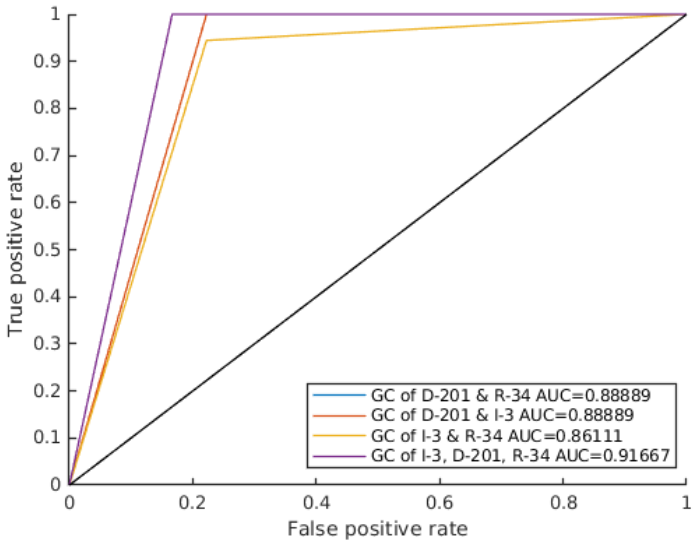


Figure 5.9: ROC curve and AUC values of BI dataset of all combinations of Copula ensemble

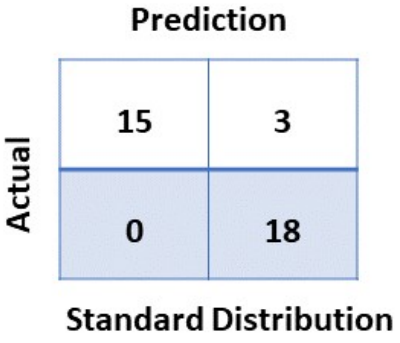


Figure 5.10: Confusion Matrix of BI dataset of the best Ensemble Model(Dense-201+Inception-V3+ResNet-34)

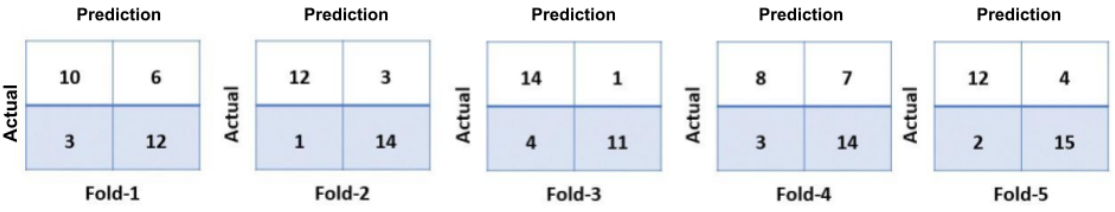


Figure 5.11: Confusion Matrices of JUCYT dataset of the best Ensemble Model(Dense-161+Inception-V3)

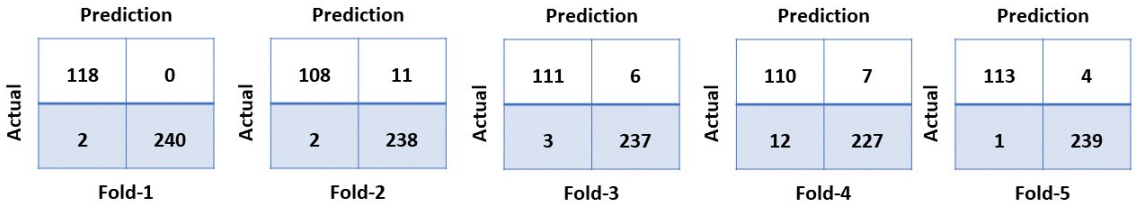


Figure 5.12: Confusion Matrices of BreakHis dataset of the best Ensemble Model(Dense-161+Inception-V3)

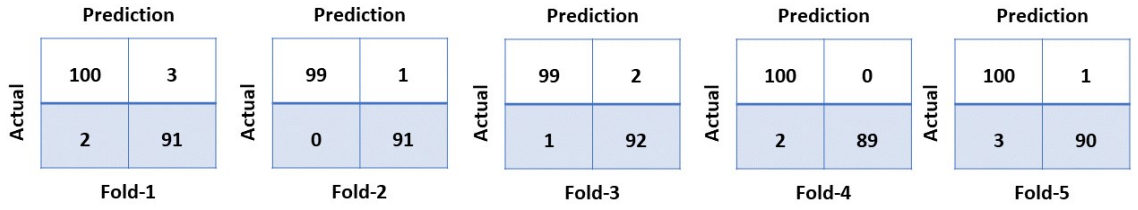


Figure 5.13: Confusion Matrices of SipakMeD dataset of the best Ensemble Model(Dense-161+Inception-V3)

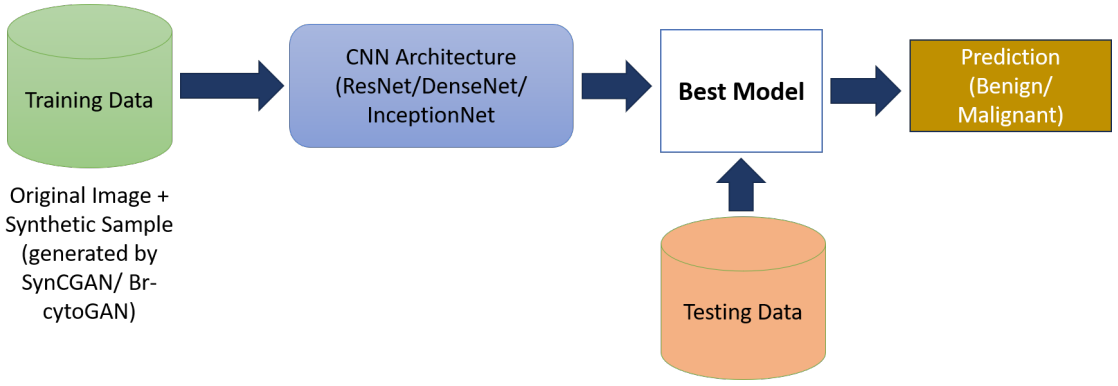


Figure 5.14: Block Diagram of Classification Process using Augmented Training set

Table 5.6: Comparative study between Gaussian Copula based Ensemble and Traditional Ensemble Methods for BI dataset (Standard Distribution)

Method	Classification Model	Accuracy	
		ROI Voting	Without ROI Voting
Individual CNN	𝕀-V3	80.56	72.22
	ℝ-34	80.56	77.78
	𝔻-201	86.11	83.33
Gaussian Copula	𝕀-V3 + ℝ-34	86.11	83.33
	𝕀-V3 + 𝔻-201	88.89	86.11
	𝔻-201 + ℝ-34	88.89	88.89
	𝕀-V3 + ℝ-34 + 𝔻-201	**91.67**	**88.89**
Average Probability	𝕀-V3 + ℝ-34	86.11	83.33
	𝕀-V3 + 𝔻-201	86.11	86.11
	𝔻-201 + ℝ-34	**88.89**	86.11
	𝕀-V3 + ℝ-34 + 𝔻-201	86.11	83.33
Majority Voting	𝕀-V3 + ℝ-34 + 𝔻-201	**88.89**	83.33
Bagging Ensemble	𝕀-V3 + ℝ-34 + 𝔻-201	86.11	

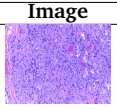
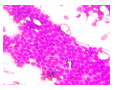
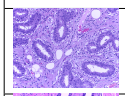
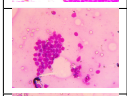
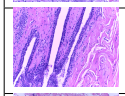
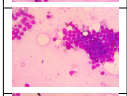
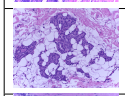
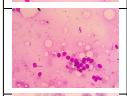
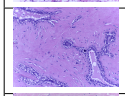
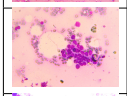
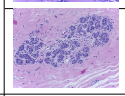
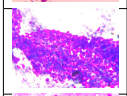
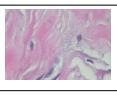
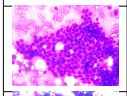
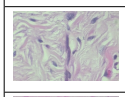
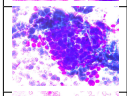
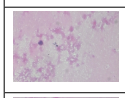
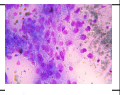
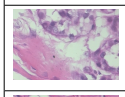
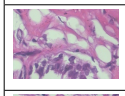
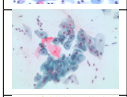
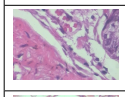
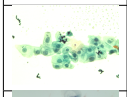
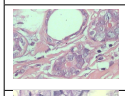
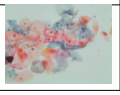
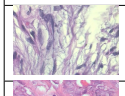
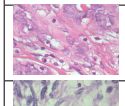
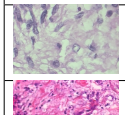
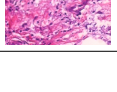
Table 5.7: Comparison Table between state of art and the proposed method

Authors Name	**Dataset**	**Classification Method**	**Accuracy(%)**
Spanhol et al. [97] (2017)	BreakHis (400X)	BVLC CaffeNet , DeCAF features	Image Level: 81.5± 2.6
Das et al. [19] (2017)		GoogLenet	Each view : 89.17
Han et al. [38] (2017)		CSDCNN	Image Level: 94.9 ± 2.8 Patient Level: 95.7±2.2
Kurmi et al. [60] (2019)		MLP	96.96
Mesut et al. [105] (2020)		BreastNet	Image Level: 95.88
Abbasniya et al. [1] (2022)		IRv2-CXL	Image Level: 96.15
Proposed Work		InceptionNet-v3 and DenseNet-161 fusion by GC	ROI based Majority voting: **97.32±1.8**
Araújo et al. [7] (2017)	BI	CNN+SVM	83.3% for binary classification
Proposed Work		InceptionNet-V3, DenseNet-201 and ResNet-34 fusion using GC	ROI based Majority voting: **91.67** for binary classification
Win et al. [114] (2020)	SipakMeD	Majority Voting based fusion of linear discriminant, SVM, KNN, Boosted trees, Bagged trees	98.27% for binary classification
Proposed Work		InceptionNet-V3 and DenseNet-161 fusion by GC	ROI based Majority voting: **98.46±0.9**(5-fold cross validation) 98.979% (3:1:1 ratio) for binary classification

Table 5.8: Precision, F1 Score, Specificity, Sensitivity analysis of JUCYT, BreakHis, SIPaKMeD and BI datasets for the best fusion model

Data Distribution	Accuracy	Sensitivity	Specificity	Precision	False Positive Rate	F1 Score	Matthews Correlation Coefficient
Gaussian Copula based Ensemble Model (DenseNet-161+ InceptionNet-v3) For JUCYT Dataset							
Fold-1	0.71	0.63	0.8	0.77	0.2	0.69	0.43
Fold-2	0.87	0.8	0.93	0.92	0.07	0.86	0.74
Fold-3	0.83	0.93	0.73	0.78	0.27	0.85	0.68
Fold-4	0.69	0.53	0.82	0.73	0.18	0.62	0.37
Fold-5	0.82	0.75	0.88	0.86	0.12	0.8	0.64
Average	0.78±0.08	0.73±0.16	0.83±0.08	0.81±0.08	0.16±0.08	0.76±0.11	0.57±0.16
Gaussian Copula based Ensemble Model (DenseNet-161+ InceptionNet-v3) For BreakHis Dataset							
Fold-1	0.99	1	0.99	0.98	0.008	0.99	0.98
Fold-2	0.96	0.91	0.99	0.98	0.008	0.94	0.92
Fold-3	0.97	0.95	0.98	0.97	0.013	0.96	0.94
Fold-4	0.94	0.94	0.94	0.9	0.05	0.92	0.88
Fold-5	0.98	0.97	0.99	0.99	0.004	0.97	0.97
Average	0.97±0.02	0.95±0.03	0.98±0.02	0.96±0.04	0.017±0.02	0.95±0.03	0.94±0.04
Gaussian Copula based Ensemble Model (DenseNet-161+ InceptionNet-3) For SIPaKMeD Dataset							
Fold-1	97.45	0.97	0.97	0.98	0.021	0.97	0.94
Fold-2	99.47	0.99	1	1	0	0.99	0.98
Fold-3	98.45	0.98	0.99	0.99	0.01	0.98	0.97
Fold-4	98.95	1	0.97	0.98	0.02	0.99	0.97
Fold-5	97.93	0.99	0.96	0.97	0.03	0.98	0.95
Average	98.45±0.72	0.99±0.01	0.98±0.012	0.98±0.011	0.017±0.012	0.98±0.008	0.96±0.01
Gaussian Copula based Ensemble Model (DenseNet-201+ InceptionNet-V3+ResNet-34) For BI Dataset							
Standard Distribution	91.67	83.33	1	1	0	90.91	84.52

Table 5.9: Some examples of cytology and histopathology images along with their classification results using individual classifiers and their ensemble techniques. Here, ✓& 0 indicates that the image has properly classified and misclassified respectively. $\mathbb{R}$= ResNet, $\mathbb{D}$= DenseNet, $\mathbb{I}$= InceptionNet-V3, GC= Gaussian Copula, MV= Majority Voting, AP= Average Probability.

Dataset	Image	$\mathbb{R}$	$\mathbb{D}$	$\mathbb{I}$	GC	MV	AP	Dataset	Image	$\mathbb{R}$	$\mathbb{D}$	$\mathbb{I}$	GC	MV	AP
BI		0	0	0	0	0	0	JUCYT		0	✓	0	✓	0	✓
		✓	0	✓	0	0	0			0	✓	✓	✓	✓	✓
		0	✓	0	✓	0	0			✓	✓	✓	✓	✓	✓
		✓	0	✓	✓	✓	✓			0	✓	0	✓	0	0
		✓	✓	0	✓	✓	✓			0	✓	0	✓	0	✓
		✓	✓	✓	✓	✓	✓			✓	✓	0	✓	✓	✓
BreakHis		0	0	✓	✓	0	0			0	✓	0	0	0	0
		✓	✓	✓	✓	✓	✓			✓	0	✓	0	0	0
		0	0	✓	0	0	0			✓	0	✓	0	✓	✓
		✓	✓	0	✓	✓	✓			✓	0	✓	✓	✓	✓
		✓	✓	✓	✓	✓	✓	SIPAKMED		✓	0	✓	✓	✓	✓
		0	✓	✓	✓	✓	✓			✓	✓	0	0	✓	0
		0	✓	✓	0	✓	✓			✓	0	✓	✓	✓	0
		✓	0	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓
		0	0	0	0	0	0								
		✓	0	0	✓	0	0								
		0	0	✓	✓	0	✓								
		0	✓	0	✓	0	0								

5.4.3 Breast Cytology Image Classification using Augmented Training Set

Data augmentation can enhance classification performance in the diagnosis of ductal carcinoma. In this study, we employed traditional CNN models to identify malignancy, training them on augmented datasets. These datasets were generated using both synthetic data generation and conventional data augmentation techniques, enabling a comparative analysis. The block diagram of classification process using augmented training set is mentioned in Figure 5.14.

For the experiments with synthetic data generated by SynCGAN model(mentioned in Chapter 3), we have used three classifiers, namely ResNet-152($\mathbb{R}-152$) [39], Inception-V3($\mathbb{I}-3$) [100] and DenseNet-161($\mathbb{D}-161$) [41]. The experiment is conducted on JUCYT breast cytology image dataset. The data description and data distribution for the train, test, and validation sets are mentioned in Table 3.1. At the time of training, the hyperparameters like batch size, number of epochs, and learning rate are set to 8, 200, and 0.001 respectively.

The outcome of these experiments will be helpful in analyzing the impact of synthetic data augmentation on several grounds.

1. Impact of data augmentation using SynCGAN on the classification pereformance.
2. Performance of SynCGAN data augmentation against traditional data augmentation.
3. Performance of SynCGAN generated data against GAN generated data.

The first observation as shown in Table 5.10, shows that augmentation of data with SynCGAN improves the performance of classifiers. For this case, DenseNet-161 has achieved 86.67% accuracy. When compared with traditional augmentation techniques like random horizontal and vertical flipping, random rotation and addition of Gaussian noise, the proposed method of augmentation has a higher impact. When traditional data augmentation was combined with SynCGAN based augmentation, the performance was either at par or lower than exclusive SynCGAN based augmentation.

We also implemented a purely GAN based pipeline for data augmentation, which performed far below the proposed model as shown in Table 5.11. This GAN based architecture also had a generator and discriminator network similar to our proposed model for a fair comparison.

Table 5.10: Performance of classifiers while using the dataset with and without augmentation. Orig: Original Data, Prop: Data generated using the proposed SynCGAN pipeline, Trad: Data generated using traditional augmentation techniques

Classifier	Orig	Orig+Prop	Orig+Trad	Orig+Trad+Prop
$\mathbb{R}-152$	73.33	**76.67**	66.66	**76.67**
$\mathbb{D}-161$	80.00	**86.67**	60	84.67
$\mathbb{I}-3$	73.33	**80.00**	63.33	76.67

Table 5.11: Performance of classifiers while using the dataset with and without augmentation. Orig : Original Data, Prop: Data generated using proposed SynCGAN pipeline, GAN: Data generated using GAN.

Classifier	Orig	Prop	Orig+Prop	GAN	Orig+GAN
$\mathbb{R}-152$	73.33	73.33	**76.67**	50.00	63.33
$\mathbb{D}-161$	80.00	63.33	**86.67**	50.00	60.00
$\mathbb{I}-3$	73.33	66.67	**80.00**	56.67	66.67

Another data augmentation technique-BrCytoGAN is implemented to generate selective synthetic high-quality cytology samples. This data augmentation technique is described in detail in Chapter 3(see Section 3.3.4). The selective synthetic samples are used as an augmented training set and the classification performances in the breast cytology domain are also analysed. Here we have used two breast cytology image datasets: 1) JUCYT dataset, and 2) the Breast cytology Pap smear dataset [89]. The detailed descriptions of these datasets are already mentioned in Chapter 2(see Section 2.3). During the classification phase, when the CNNs are trained using the augmented dataset by BrCytoGAN, the hyperparameters like batch size, number of epoch, optimizer, learning rate are fixed as determined by the SynCGAN model.

As a selection model, fuzzy entropy based ensemble rule is explored. Here, we have used traditional CNN models like ResNet-18($\mathbb{R}-18$), DenseNet-161($\mathbb{D}-161$), and InceptionNet-v3($\mathbb{I}-3$) as the base model for the ensemble. The different combinations of CNNs are ensembled for the selection procedure. Table 5.12 shows that the selection process by multiple CNNs is performing better than the individual CNN. Also, it is found that the classification performance on the test set for DenseNet-169 model and InceptionNet-V3 both has achieved 96.55% accuracy for the selection of multiples CNNs ensemble and entropy-based rule. In Table

5.13, we present a comparison of classification performance using various training sets(details data distribution is mentioned in Table 3.2), including the original set, an augmented set with selection, and an augmented set without selection. The Densenet-169 model achieved an accuracy of 93.10% when trained with the augmented set that included selection, compared to an accuracy of 86.21% obtained with the original training set. In addition, we have achieved an accuracy of 82.76% in the DenseNet-169($\mathbb{D}-169$) model, trained on the augmented train set (without selection), which performs 10.34% lower than the model trained on selective synthetic images. Also, we have constructed another train set after the first phase of the selection process, i.e., the images are selected by the joint performances of the CNNs. Here, we have selected only those synthetic samples, which have the prediction probability of winning the class is higher. But in that case, we have got less accuracy from the proposed selection method for each CNN classifier.

We have experimented another selection process, where the class-specific probability distribution values of synthetic samples are calculated with respect to InceptionNet-v3 model, which is trained on original cytology samples. If the synthetic sample is predicted true with a high probability value (greater than 0.9), then the sample is selected. By this technique, we have achieved 79.31% accuracy on DenseNet-169 model, which is less than our proposed two-step selection process.

Table 5.12: Comparison of classification Performances of the test set(of JUCYT dataset) on the selection of different training sets (individual to different ensemble model)

CNN Model	Selection by Single CNN and Entropy			Selection by Multiple CNN and Entropy			
	$\mathbb{R}-18$	$\mathbb{D}-161$	$\mathbb{I}-3$	$\mathbb{D}-161$ + $\mathbb{R}-18$	$\mathbb{D}-161$+ $\mathbb{I}-3$	$\mathbb{I}-3$ +$\mathbb{R}-18$	$\mathbb{I}-3$+$\mathbb{R}-18$+$\mathbb{D}-161$
$\mathbb{D}-169$	89.65	93.10	93.10	93.10	**96.55**	**96.55**	93.10
$\mathbb{R}-18$	79.31	86.20	82.75	**93.10**	89.66	75.86	89.65
$\mathbb{I}-3$	93.10	93.10	89.66	93.10	93.10	**96.55**	89.66

Also, we have compared with another traditional ensemble rule like average probability in place of fuzzy rank-based voting in the first phase of the proposed selection rule, where, we have achieved 82.76% accuracy on DenseNet-169 model, which is less than the proposed method. We have also created another train set by some traditional data augmentation techniques like flipping (horizontal and vertical), random crop, rotation at a 30-degree, adding noise, etc., where, we have achieved 79.31% accuracy on DenseNet-169 model. Another Pap Stain-based breast cytology image dataset (Sakia et al.) is also used to validate the proposed BrCytoGan model. Table 5.14 presents the classification performance of various CNN models using different training sets selected by either a single classifier or an ensemble of classifiers. Among them, the fuzzy ensemble of DenseNet-169,

Table 5.13: Classification performance on the test set(of JUCYT dataset), trained by different selections of the training set. (Selection by the ensemble of three CNN models- $\mathbb{I}-3$, $\mathbb{R}-18$, $\mathbb{D}-161$)

Classification Model	Original Train Set	Augmented Train Set (without selection)	Augmented Train Set (with proposed selection)	Augmented Train Set (with InceptionNet-V3 based features selection)	Augmented Train Set (with fuzzy ensemble based selection)	Augmented Train Set (with Average probability based ensemble selection)	Augmented Train Set (Traditional Data Augmentation)
$\mathbb{D}-169$	86.21	82.76	93.10	79.31	72.41	82.76	79.31
$\mathbb{R}-18$	68.97	72.41	89.66	72.41	79.31	82.76	75.86
$\mathbb{I}-3$	72.4	75.86	89.66	79.31	82.76	86.21	68.97

ResNet-18, and Inception-v3, along with entropy-based selection, achieved an accuracy of 97.67% for this dataset, with Inception-v3 and DenseNet-169 contributing significantly to this performance. Also we have compared the proposed selection with other selection rules and evaluate the classification performance in each cases. The details are mentioned in Table 5.15.

In Figure 5.15 and Figure 5.16, we have studied ROC curves (True positive rate vs. False positive rate) and also the AUC (Area Under Curve) values of different classification models on the proposed augmented train set for JUCYT and breast cytology pap Stain dataset respectively. The confusion matrices of different classification models are shown in Figure 5.17 and Figure 5.18, for JUCYT and pap Stain dataset respectively. Tables 5.16 and 5.17 present an analysis of performance metrics such as Precision, Recall, and F1 score for CNN models trained on the proposed augmented dataset for both cytology datasets, respectively.

Also, a comparative study has been presented with the state-of-the-art methods of breast cytology pap Stain dataset [89] in Table 5.18.

Table 5.14: Comparison of classification Performances of breast cytology pap Stain dataset(Saikia et al.) on the selection of different training sets (individual to different ensemble model)

CNN Model	Selection by Single CNN and Entropy			Selection by Multiple CNN and Entropy			
	$\mathbb{R}-18$	$\mathbb{D}-161$	$\mathbb{I}-3$	$\mathbb{D}-161$+$\mathbb{R}-18$	$\mathbb{D}-161$+$\mathbb{I}-3$	$\mathbb{I}-3$+$\mathbb{R}-18$	$\mathbb{I}-3$+$\mathbb{R}-18$+$\mathbb{D}-161$
$\mathbb{D}-169$	95.35	95.35	93.02	97.67	93.02	93.02	**97.67**
$\mathbb{R}-50$	86.05	88.37	88.37	93.02	81.40	93.02	**93.02**
$\mathbb{I}-3$	95.35	93.02	95.35	93.02	86.05	97.67	**97.67**

5.4.4 Classification Model using feature optimization technique

Optimal features selection is an utmost importance to enhance the performance of cytology image classification. In this work, we have proposed Artificial Electric

Table 5.15: Classification performance on breast cytology pap Stain dataset(Saikia et al.), trained by different selections of the training set. (Selection by the ensemble of three CNN models- $\mathbb{I}-3$, $\mathbb{R}-18$, $\mathbb{D}-161$)

Classification Model	Original Train Set	Augmented Train Set (without selection)	Augmented Train Set (with proposed selection)	Augmented Train Set (with InceptionNet-V3 based features selection)	Augmented Train Set (with fuzzy ensemble based selection)	Augmented Train Set (with Average probability based ensemble selection)	Augmented Train Set (Traditional Data Augmentation)
$\mathbb{D}-169$	93.02	90.69	**97.67**	86.04	93.02	93.02	95.34
$\mathbb{R}-50$	90.70	90.70	**93.02**	83.72	90.70	90.70	90.70
$\mathbb{I}-3$	95.35	95.35	**97.67**	95.35	95.35	93.02	95.35

Table 5.16: Class wise Precision, recall, F1 scores of different CNNs(trained on augem- nted set of proposed selection model-Ensembled of three CNNs + Entropy) on Test set(JUCYT dataset).

CNN Model	Precision		Recall		F1-score	
	Benign	Malignant	Benign	Malignant	Benign	Malignant
$\mathbb{D}-169$	1	0.91	0.78	1	0.88	0.95
$\mathbb{R}-18$	0.88	0.9	0.78	0.95	0.82	0.93
$\mathbb{I}-3$	0.88	0.9	0.78	0.95	0.82	0.93

Table 5.17: Class wise Precision, recall, F1 scores of different CNNs(trained on augemnted set of proposed selection model-Ensembled of three CNNs + Entropy) on Test set(Saikia et al.).

CNN Model	Precision		Recall		F1-score	
	Benign	Malignant	Benign	Malignant	Benign	Malignant
$\mathbb{D}-169$	0.95	1	1	0.95	0.97	0.97
$\mathbb{R}-50$	0.94	0.91	0.9	0.95	0.92	0.93
$\mathbb{I}-3$	1	0.95	0.95	1	0.97	0.97

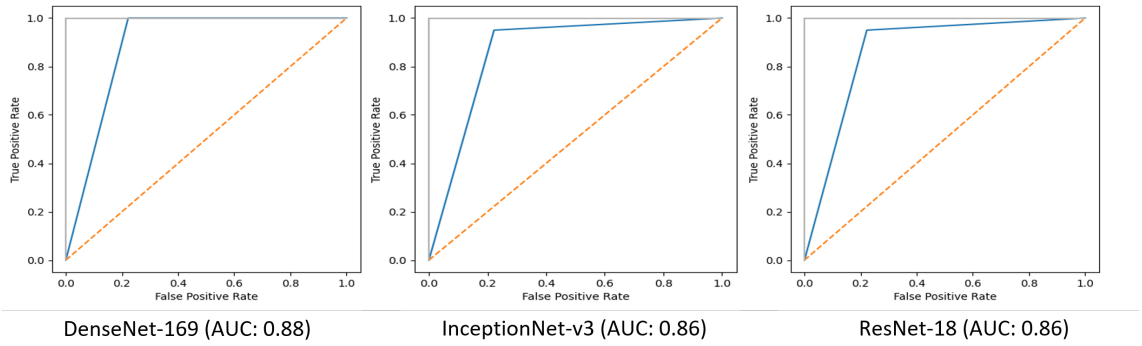


Figure 5.15: ROC curves and AUC values of different CNN models, which are trained on Original & selective synthetic cytology image set(Ensemble by three base classifiers)-JUCYT dataset

Field Algorithm (AEFA) [42] to find the optimally appropriate features by discarding the less relevant ones in order to boost the performance of cytology image classification. Here, the features are first extracted from the traditional ResNet18 [39] model. Finally, this optimal subset of features are classified by a SVM (Support Vector Machine) classifier. The details are described in the following paragraphs:

Proposed Methodology:

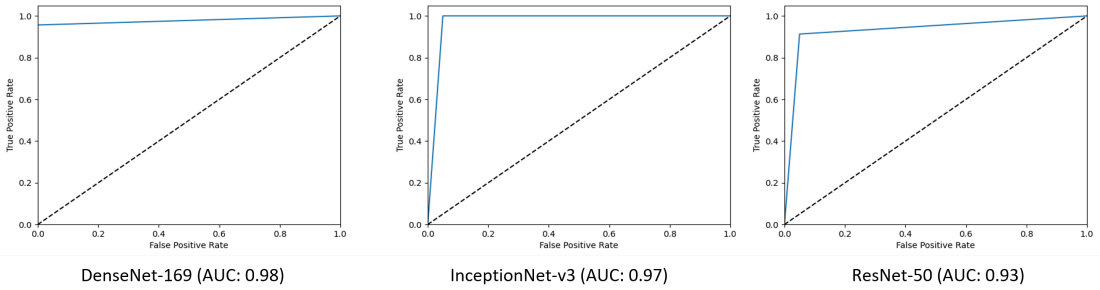


Figure 5.16: ROC curves and AUC values of different CNN models trained on Original & selective synthetic cytology image set (Ensemble by three base classifiers) of breast cytology pap Stain dataset(Sakia et al.)

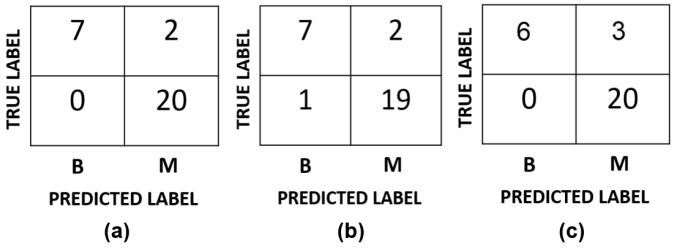


Figure 5.17: Confusion Matrices of different classifiers, which are trained on Original and selected(Ensemble by three base classifiers) synthetic cytology image set(a: DenseNet-169, b: InceptionNet-v3, c: ResNet-18)-JUCYT dataset

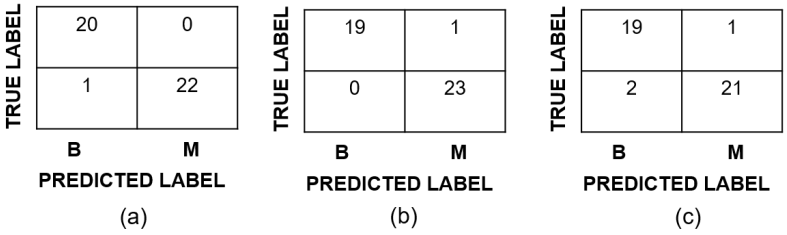


Figure 5.18: Confusion Matrices of different classifiers, which are trained on Original and selected(Ensemble by three base classifiers) synthetic cytology image set(a: DenseNet-169, b: InceptionNet-v3, c: ResNet-50)-Breast cytology Pap Stain dataset(Sakia et al.)

The Artificial Electric Field Algorithm (AEFA)[42] is a metaheuristic optimization algorithm primarily based on Coulomb's law of electrostatics. Coulomb's law states that the electrostatic force (F) of attraction or repulsion between two charged particles(Q_1 and Q_2) is proportional to the product of the magnitude of their charges and inversely proportional to the square of the distance(R) between them, i.e. $F = K\frac{Q_1 Q_@}{R^2}$,; where K is Coulomb's constatnt. The electrostatic force repels or attracts the charged particles in search space. Let's consider the search space of an optimization problem to be the space and the population of the solution to

Table 5.18: Comparison of the classification accuracies of the proposed approach with the state-of-the-art method for breast cytology pap Stain dataset(Saikia et al.)

Author	Methodology	Result(in %)
Saikia et al. (2019)	GoogLeNet V3 fine tuned(Traditional data Augmentation)	96.25
Proposed	$\mathbb{I} - 3$(Augmentation by BrCyto-GAN)	**97.67**

be the charged particles residing in them; we can see that the particles change position due to the electrostatic forces among them. Also, the electrostatic force is the only connecting link among the charges, and as a result, the problem solution takes the position of the charges in the space. In AEFA, the magnitude of charges of the particles is considered a fitness factor to evaluate the population, and the positions of the particles in the search space are considered as the solution to the optimization problem. The AEFA is seen as a closed system of charges that obeys Coulomb's law of electrostatic force and Newton's laws of motion.

The challenge of selecting the best optimal subset of features is NP-hard and requires an exhaustive search. With an increment in the number of features, the state space of the search algorithm gets expanded exponentially. To overcome this, we have used AEFA to select the optimal subset of features from the given set of features. Feature selection using the optimization algorithm generally works in the following way: (1) From the given feature space, a subset of features is selected based on the optimization algorithm. (2) The fitness value of the selected features is calculated. (3) A new features subset is generated using the optimization algorithm. (4) The previous three steps are continually executed until the termination condition is reached. The flowchart of the proposed methodology is shown at Figure 5.19.

The optimal feature selection algorithm's goal is to choose a significantly better-fitted feature set of dimension c from the initial feature set of dimension d, such that $c \leq d$. Firstly, a random set of the population of N particles $S = [S_1, S_2, ..., S_N]^T$ is sampled. Here, every $S_i = [s_{i,1}, s_{i,2}, ..., s_{i,d}]$ represents a binary vector of a subset of features in a d dimensional feature space. $s_{i,d} = 0$ represents the fact that the feature is being discarded, and $s_{i,d} = 1$ signifies that the feature is being present or, in other words, selected.

The classification method considers the randomly selected features for fitness (accuracy) computation. The accuracy or the fitness of the sample is calculated using the Support Vector Machine with RBF kernel. The optimal feature selection

method aims to choose the substantially appropriate subset of optimal attributes to maximize the fitness value. The workings of the AEFA for the i^{th} sample of the population is shown at Figure 5.20.

The optimal features selection technique is described in the following algorithm (see Algorithm 5):

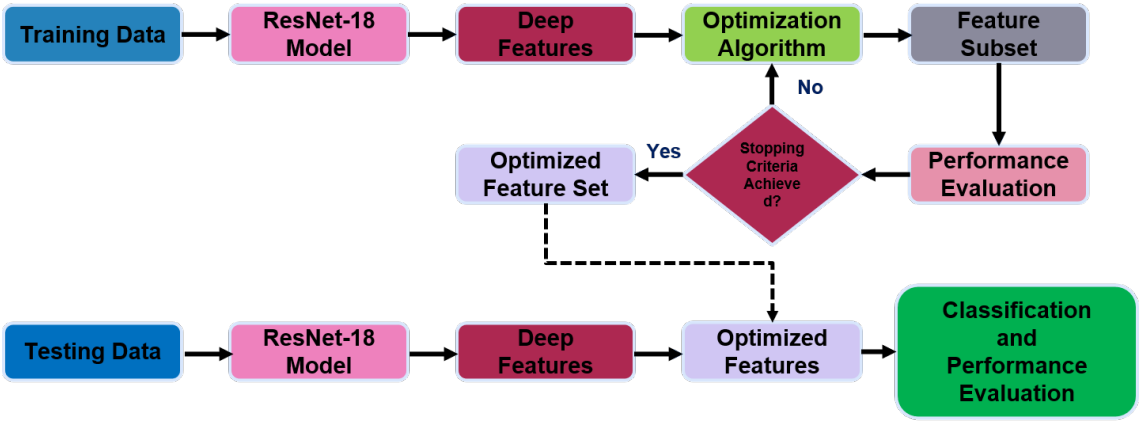


Figure 5.19: Flow Diagram of Proposed Malignancy Identification Technique using Optimal Feature Selection

Experimental Setup:

For this methodology, we have used JUCYT breast cytology image dataset, which consist of 203 cytology samples. Among these samples, 94 are benign and 109 are malignant samples. The dataset descriptions in details are mentioned in Chapter 2. It is split into 75%, 13%, and 12% for train, test and validation sets.

We employed the ResNet-18 architecture, pretrained on the ImageNet dataset with 1000 classes, as a feature extractor CNN. The model was trained in a PyTorch environment and executed on an NVIDIA-510 GPU with 6GB memory. During training, the hyperparameters were set as follows: batch size of 4, 200 epochs, Adam optimizer, and Negative Log-Likelihood (NLL) loss function. Features were extracted from the convolutional layer preceding the fully connected layer of the best-performing ResNet-18 model.

Result & Discussion:

The test set is classified using a Support Vector Machine (SVM) classifier with a Radial Basis Function (RBF) kernel. The maximum number of iterations(MaxIteration), population size(N), feature space dimension(d) are set to 10, 25, 512 respectively. The initial Coulomb constant(K_0) value is chosen as 500.

Algorithm 5: Optimal Features Selection Technique by AEFA

The values of d (dimension), N (total number of subsets) and **MaxIteration** (maximum number of iterations) are initialized.
The binary feature subsets $(S_1, S_2, S_3, ..., S_N)$ are initialized randomly(i.e. initialization of population of size N) in the search range $[S_{min}, S_{max}]$
The velocities are initialized to a random value.
The fitness values $(f_{S_1}, f_{S_2}, f_{S_3}, ..., f_{S_N})$ of $(S_1, S_2, S_3, ..., S_N)$ are evaluated. Here, the fitness values are the recognition accuracies, made by the SVM classifier.
Set iteration at $g = 0$
while Stopping Criteria is not achieved **do**
    Coulomb's constant (K_g), best fitness($best(g)$) and worst fitness ($worst(g)$)are calculated as follow:
    $K_g = K_0 * exp(-\alpha \frac{g}{MaxIteration})$,
    where α and K_0 decay rate and initial Coulomb constant respectively.
    $best(g) = max(f_{S_j})$ and $worst(g) = min(f_{S_j}); j \in (1, 2, ..N)$
    for $i \leftarrow 1$ to N **do**
        Fitness value f_{S_i} is evaluated
        Calculate charge q_i on the ith particle by $q_i = exp(\frac{(f_{S_i} - worst(g))}{(best(g) - worst(g))})$
        $Q_i = \frac{q_i}{\sum_{i=1}^{N} q_i}$
        The total force $F_{i,d}^g$, the total electric field $e_{i,d}^g$ and acceleration $a_{i,d}^g$ is calculated by following formulas:
        $F_{i,d}^g = \sum_{j=1, i \neq j}^{N} rand() F_{ij}^d(g)$; where $F_{ij}^d(g)$ is the force acting on the charge i from charge j. $e_{i,d}^g = \frac{F_{i,d}^g}{Q_i}$ $a_{i,d}^g = \frac{e_{i,d}^g * Q_i}{M_i(g)}$; $M_i(g)$ is the unit mass of ith particle.
        $v_{i,d}^{g+1} = rand() v_{i,d}^g + a_{i,d}^g$ and $S_{i,d}^{g+1} = S_{i,d}^g + v_{i,d}^{g+1}$
    end for
    Probability is calculated for each dimension velocity, and the binary encoding function is used, which creates the binary vector from the feature space.
        $prob(S_{i,d}^{g+1}) = \frac{1}{1 + exp(-S_{i,d}^{g+1})}$
        if $prob(S_{i,d}^{g+1}) > 0.5$ **then** $S_{i,d}^{g+1} = 1$
        else $S_{i,d}^{g+1} = 0$
    New population is formed after $S_{i,d}^{g+1}$. New population is the set of updated positions of all particles.
end while
Finally, the optimal subset is returned.
The accuracy for the optimal subset is checked against validation data.

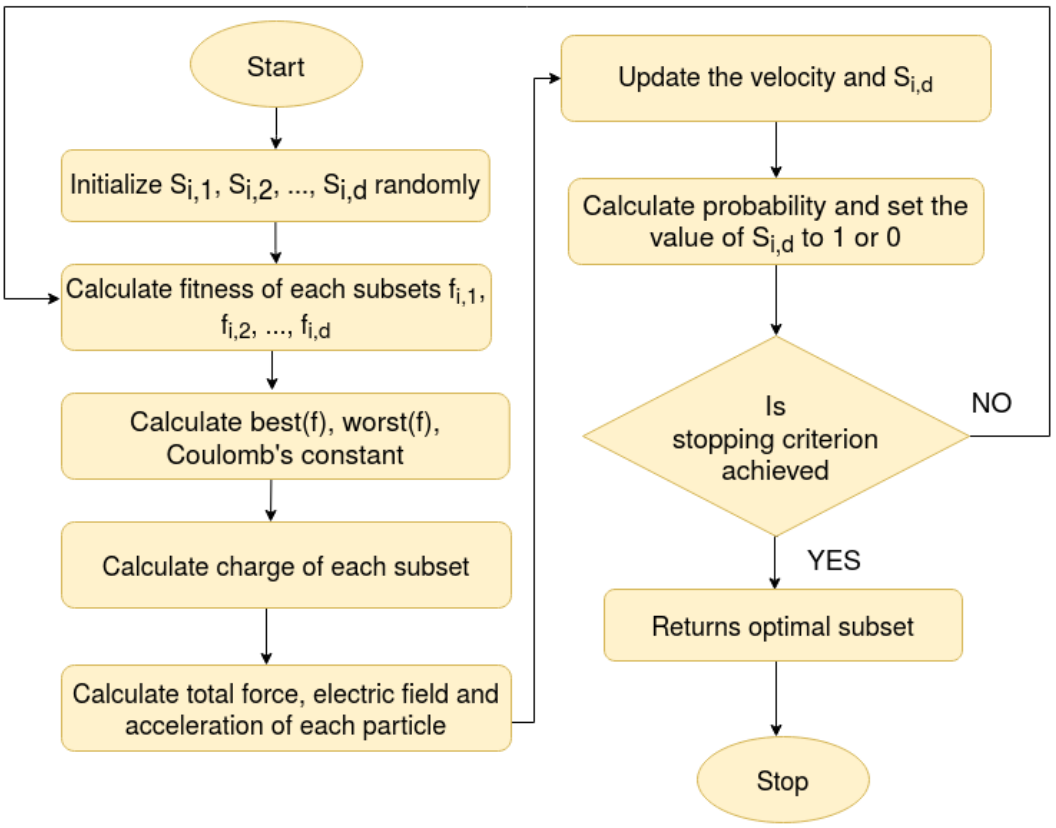


Figure 5.20: Flow Diagram of Optimal Features selection using AEFA

Table 5.19: Performances of SVM classifier by with and without optimal features selection

Performance Metrices	Without Feature Selection	With Feature Selection
Number of features	512	91
Accuracy(%)	0.62	**0.85**
F1 Score	0.61	0.83
Precision	0.62	0.81
Recall	0.64	0.82

The confusion matrices before and after feature selection are demonstrated in Figure 5.21.

By using the complete set of 512 features, the SVM model achieved an accuracy of 62%, precision of 62%, recall of 64%, and an F1 score of 61%. After selecting the optimal subset of 91 features using the Artificial Electric Field Algorithm (AEFA), the performance improved significantly, achieving an accuracy of 85%, precision of 71%, recall of 82%, and F1 score of 83%. This demonstrates that the optimal feature selection process outperformed the result in cytology image classification.

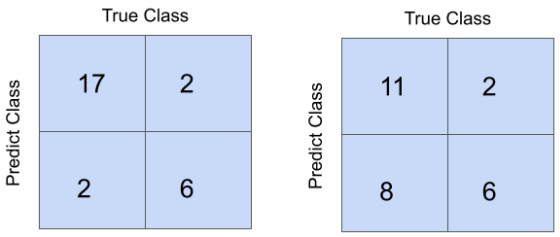


Figure 5.21: Confusion matrices of classification by SVM. Left: After Optimal feature selection, Right: Before features selection

A detailed comparison of the performance with and without feature selection is provided in Table 5.19.

5.4.5 Classification performances of Different Methodologies on the Complete JUCYT dataset

In this thesis, the previously discussed proposed methodologies have been evaluated on different subsets of the JUCYT dataset (with dataset sizes of 150, 156, 203, and 212 for the SynCGAN, GC-EnC, AEFA-based features optimization, and BrCyto-GAN models, respectively), as the volume of the JUCYT dataset has increased over the year. To ensure a fair comparison and provide standardized results across all models based on a fixed data distribution, we have re-evaluated all the models using the 212-sample breast cytology image dataset. The detailed data distribution is provided in Table 3.2. Here, we have reported the classification performance only for the best-performing models previously identified for each proposed methodology. For GC-EnC, the best-performing model is the Gaussian Copula ensemble of InceptionNet-v3 and DenseNet-161 on the JUCYT dataset. For the data augmentation models like SynCGAN and BrCyto-GAN, the best models are found to be DenseNet-161 and InceptionNet-v3(or DenseNet-169), respectively. The experimental results for the best models on the complete dataset are mentioned in Table 5.20. The GC-EnC based ensemble model has achieved a classification accuracy of 86.2% when experimented on the complete JUCYT dataset. The DenseNet-161 has performed classification accuracy of 89.65% when trained on augmented set by SynCGAN model(synthetic samples are generated using the training set of the total JUCYT dataset). For this data augmentation model, we have primarily generated 4000 synthetic samples, and from these samples, 180 samples are chosen by visual inspection. Training the CNNs on the augmented set

by BrCyto-GAN model, we have found the highest classification performance, i.e. 96.55% accuracy for both the InceptionNet-v3 and DenseNet-169 models. We have incorporated this trained DenseNet-169 model into the web server for automated malignancy identification in real time.

5.4.6 Web application for breast cytology image classification

In this study, we have created a Web application designed for the automated classification of breast cytology images within a real-time setting. In this application, the best-trained model of DenseNet-169, which is trained by augmented dataset of breast cytology images generated by BrCyto-GAN model(steps are mentioned in Chapter 3.3.4), is used for classification. The basic system software and software tools required for running the web application are as follows:

1. Operating System: Windows 7 or later, / Linux: Ubuntu 16.04 or later/ Mac OS 10.12.6 or later.
2. Python 3.7 or later
3. Flask
4. Pytorch
5. Numpy
6. Matplotlib

Steps to use the WEB Application:

For using this application, one has to send the http request to the url:" `http://nibarancse.jdvu.ac.in/projects/cytology/`" and a web page will be opened. Then the breast cytology image (image format: .jpg, .png) is to be uploaded by clicking the specific button of the web page. After that this uploaded image will be tested in our web server, and the classification result will be displayed within a few seconds(minimum 5 seconds). In Figure 5.22 we have described how the user will use this web application.

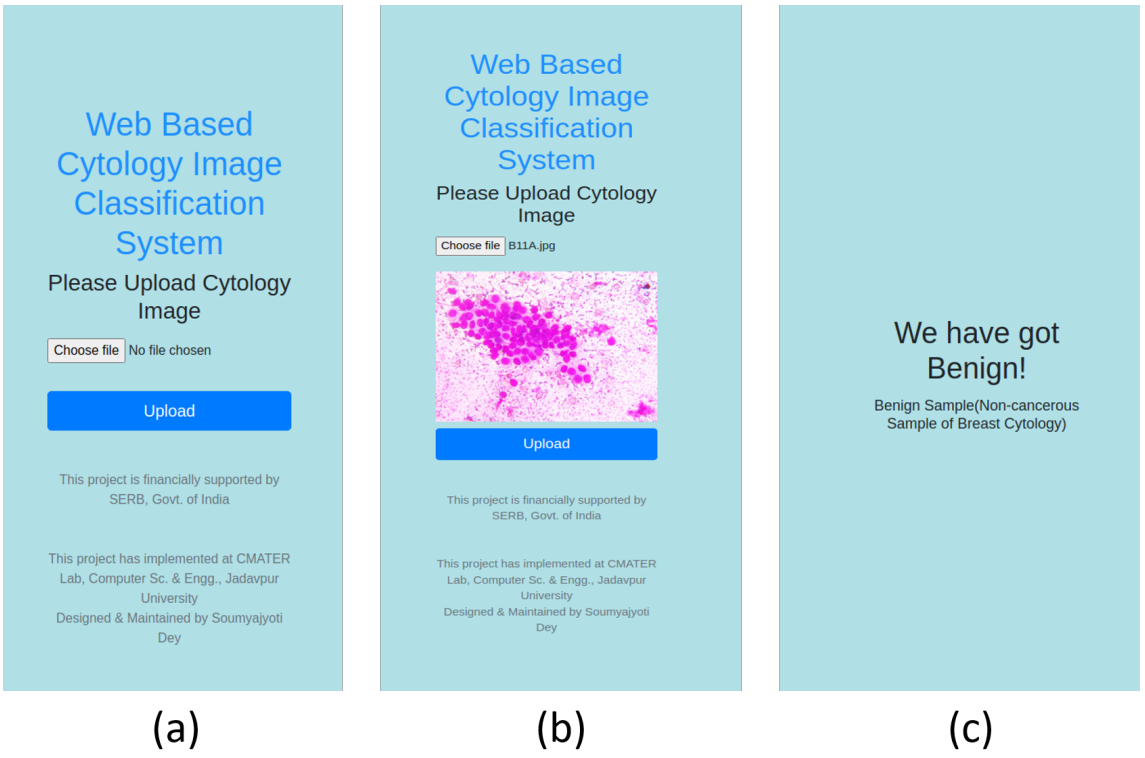


Figure 5.22: Web application (GuI) for Cytology Image Classification (a) Opening Page of the web application, (b) Upload digital cytology image, (c) Predicted result page of the application

5.5 Discussions

This chapter discusses various techniques for the automated diagnosis of malignancy from breast cytology images. Primarily, traditional CNN architectures, including different versions of ResNet, DenseNet, and InceptionNet, are used for classification. We introduce a Gaussian Copula-based ensemble rule for malignancy identification in breast cytology, which demonstrates promising results by combining Inception-v3 and DenseNet-161 for the automated diagnosis of ductal carcinoma. Beyond breast cytology, we have also tested this technique on breast histopathology and cervical cytology datasets to validate its effectiveness. Additionally, we explore the artificial electric field optimization algorithm for classifying breast cytology samples. By selecting optimized features, this method enhances performance in an SVM classifier. Furthermore, we analyze the impact of data augmentation on breast cytology classification, introducing previously established augmentation techniques and comparing different augmented training sets. Additionally, breast cytology images are classified using feature extraction

Table 5.20: Comparative study of different proposed models on JUCYT dataset

Model Name	Dataset	Methodology	Classification Accuracy in %
SynCGAN	**JUCYT (Total Samples=212) Train=156 Test = 29 Validation = 27**	CNNs are trained by the augmented training set. Augmentation is done by hybridization of GAN, CGAN models. Synthetic samples are selected by visual interpretation.	DenseNet-161:89.65
GC-EnC		CNNs are trained by the ROIs of original cytology images, extracted by fuzzy membership function. Finally samples are classified by Gaussian Copula ensemble and ROI voting rule	Gaussian Copual Ensemble of InceptionNet-v3 and DenseNet-161 with ROI voting approach: 86.20
AEFA		Optimized features are selected (100 numbers of features are selected) by Artificial Electric Field Algorithm and It is classified by SVM	SVM(RBF): 85.18
BrCyto-GAN		CNNs are trained by the augmented training set. Augmentation is done by hybridization of GAN, SRGAN and CGAN models. Synthetic samples are selected through the fuzzy based ensemble of multiple CNNs and entropy rule.	InceptionNet-V3/ DenseNet-169: **96.55** Selection of synthetic sample by fuzzy ensemble of InceptionNet-v3 and ResNet-18 & Entropy

from segmented binary masks(mentioned in the previous chapter), employing traditional machine learning classifiers. Finally, we have developed a web-based application to enable remote classification of cytology samples.

Chapter 6

Conclusion

This thesis has successfully addressed some challenges of the automated diagnosis of ductal carcinoma from breast cytology images. The major contributions of this thesis include the development of a new breast cytology image database, the proposal of novel data augmentation techniques, the implementation of models for cellular object segmentation, and the classification of breast cytology images into benign and malignant categories. Additionally, a web application has been deployed to enable real-time automated diagnosis of breast cytology images.

Development of new breast cytology image dataset—**JUCYT** is one of the major contributions of this thesis, which contains both cancerous and non-cancerous samples collected from Indian patients. The dataset comprises 212 cytology images, each with a resolution of 960×1280 pixels captured at 40x magnification. Details of the data collection and annotation process are also discussed in Chapter 2 of this thesis. Different subsets of the JUCYT dataset have been used for the experimental purposes. Due to the limited availability of publicly accessible breast cytology datasets, additional digital pathology samples, such as breast histopathology and cervical cytology images, have been utilized to validate the proposed models.

Due to the high cost associated with data collection and annotation, both deep learning-based and traditional data augmentation techniques have been discussed in Chapter 3. To mitigate these major hardel a fuzzy template based data augmentation technique and synthetic data generation models like SynCGAN and Br-CytoGAN model have been proposed. By using the selected ROIs by fuzzy

template based rule from the breast cytology samples, we have got 71.87% accuracy on InceptionNet-v3 model(the numbers of training and testing samples are 93 and 31 respectively), which is 9.37% higher than that obtained through traditional data augmentation methods. Also, after using the synthetic samples at training time, generated through SynCGAN model, the InceptionNet-v3 (90 and 30 cytology images are chosen as the training and testing sets) has achieved 80% accuracy, which is 16.67% higher than that achieved by traditional data augmentation techniques. However, in the DenseNet-161 model (trained by augmented set of SynCGAN mode), we have got the maximum classification accuracy 86.67%. However, the generated samples through SynCGAN, are of lower resolution and selected through visual inspection and expert validation. To overcome this limitation, a hybrid generative model named Br-CytoGAN is introduced. This model generates high-resolution, selectively synthetic samples. This synthetic sample selection process is guided by the joint performances of popular CNN architectures like DenseNet-169, InceptionNet-v3 and ResNet-18. After using the synthetic samples by Br-CytoGAN model, we have got classification accuracy 96.55% by the InceptionNet-v3 model(here, 156 images are taken for training and 29 images are used for testing), which is 27.58% higher than the classification accuracy achieved through the traditional data augmentation technique on the JUCYT dataset. The quality of synthetic samples has been measured by the well-known metrics such as IS, KID, FID score. Finally, it is proved that the synthetic selective samples are of better quality than SynCGAN and other state-of art data generation techniques. Additionally, the feasibility of generating synthetic breast cytology images from breast histopathology images has been investigated as part of an unpaired domain-to-domain image translation task. Two popular models—CycleGAN and Neural Style Transfer—were explored. The experimental results demonstrate that CycleGAN-generated images exhibit greater feature similarity to actual breast cytology images compared to those generated via Neural Style Transfer.

Segmentation of cellular objects from breast cytology images posed a considerable challenge. In Chapter 4, novel deep learning based models have been presented for the segmentation of breast cytology images. Here, the performances of traditional semantic segmentation models are analyzed, and to boost the segmentation performances, ensemble rules by multiple semantic segmentation models are also explored on the JUCYT dataset. A fuzzy-rule-based ensemble of SegNet

and UNet has achieved an IoU score of 83.79%, outperforming the individual performance of UNet by 5.96%. Furthermore, a novel localization-to-segmentation framework is proposed in the combination of Faster-RCNN and MedSAM model, which showed improved segmentation performance over conventional semantic segmentation models.

Chapter 5 depicts some deep learning based models for melignancy identification from breast cytology image samples. Here, traditional CNN models like different architectures for ResNet, DenseNet, InceptionNet are fine-tuned for the classification of breast cytology images. Sometimes, to enhance classification performance of the breast cytology dataset a Gaussian copula based framework for ensemble of CNNs is explored. Experimental results indicate that the complex nature of the medical images can be efficiently addressed using the proposed model. For the JUCYT dataset, the classification performance has achieved 84.37% by the proposed ensemble of DenseNet-161 and InceptionNet-v3 model, whereas using the single CNN architecture like inceptionNet-v3 model, it is found classification accuracy of 71.87% (here previously mentioned data distribution, like 93 images for training and 31 images for testing is used). The proposed ensemble technique has outperformed other traditional ensemble rules like majority voting, bagging, and average probability. It is also applied on other digital pathology domains(like breast histopathology and cervical cytology), and has performed better than other stae-of art techniques on the corresponding domains. For classification performance improvement of the JUCYT database, we have also explored Artificial Electric Field algorithm, to select the optimal features set(features are extracted by pre-trained ResNet18 model) and finally it is classified by SVM classifier. Additionally, we have investigated the classification performance of JUCYT dataset by using traditional machine learning classifiers. Here, the handcrafted features like area, perimeter, compactness, etc. are extracted from the segmented images(discussed in the Chapter 4) for learn the models. But the deep learning-based models are outperforming than the machine learning based approaches.

The proposed methodologies, previously discussed, were evaluated on different subsets of the JUCYT dataset due to its continuously increasing volume. For standardized reporting, we have also applied these methodologies to the complete

set of 212 cytology samples. So, on InceptionNet-V3 we have achieved the highest classification accuracy 96.55%, which is trained on augmented dataset by Br-CytoGAN model. Also other techniques like GC-EnC and SynCGAN data augmentation based classification model, we have achieved accuracy 86.20% and 89.65% respectively.

In this thesis finally a web application (URL: `http://nibarancse.jdvu.ac.in/projects/cytology/`) is developed for automated diagnosis of breast cytology image. This web application will serve as a useful guide to the pathologist in delivering the appropriate diagnostic decision in real time with reduced effort and time.

Future Scopes:

Building on the focused objectives and principal accomplishments of the thesis, this section presents a potential roadmap for future developments. The following proposals may be pursued as directions for future work.

In the future, we expect to increase the volume of the JUCYT database by collecting more breast cytology samples from an Indian perspective. Additionally, we plan to incorporate samples with diverse staining variations into the dataset. For this thesis, we have only explored 40x magnification samples; in the future, we aim to include samples with lower magnifications such as 4x and 10x, and attempt to establish correlations between lower and higher magnifications. The JUCYT dataset contains staining variations among the cytology samples. Even though the samples are collected from different pathology laboratories, differences in staining preparation result in significant color variations. These discrepancies can potentially lead to an increased risk of false negatives or false positives in classification tasks. Therefore, in the future, we plan to explore deep learning-based stain normalization techniques to improve classification performance.

One of the major objectives was to develop a high-quality synthetic data generation model capable of producing synthetic cytology samples with improved IS, FID or KID scores. In the current implementation of the hybrid generative models, we have generated images with dimensions of either 256×256 or 512×512 pixels(explained in Chapter 3). Moving forward, we intend to generate synthetic cytology samples at the original image resolution. Additionally, we plan to fine-tune the generative model to enable sample generation directly from noise. In

future, different deep learning based data generation models combined with the proposed ROI generation technique may be explored.

In the previously discussed semantic segmentation models(discussed in Chapter 4), one of the major limitations was that all cellular objects were not accurately segmented. When using the MedSAM architecture for this task, its performance heavily depends on the prompts generated by the localization models. To enhance this process, we propose integrating advanced object detection models such as YOLOv8 or RetinaNet. Since we are dealing with cellular structures, the presence of significant overlapping regions in the samples makes accurate separation particularly challenging. In the future, we plan to explore geometric functions [23] and study the topology of cell structures to better address these issues. Pixel-level annotation of digital pathology samples is both time-consuming and expensive. However, deep learning-based techniques require a substantial number of annotated images for effective training. To address this, we aim to explore automated cytology image annotation methods, including semi-supervised and weakly supervised learning approaches.

In the future, we plan to incorporate additional mathematical or probabilistic models as ensemble strategies to further enhance classification performance. Explainable AI (XAI) has emerged as a significant trend in computer vision research, particularly within the medical domain. In the context of cytology, XAI can help identify which regions of an image influence the malignancy decision and highlight the most critical features for breast cancer diagnosis, thereby enabling a more detailed analysis of malignancy. Several established XAI techniques already exist, such as Saliency Maps, Grad-CAM, SHapley Additive exPlanations (SHAP), and Local Interpretable Model-Agnostic Explanations (LIME). In future work, we aim to apply these models to the JUCYT dataset and explore the development of modified or improved versions tailored to the cytology imaging context.

In this thesis, we have focused exclusively on digital pathology samples for malignancy identification. In the future, we plan to explore other imaging modalities such as MRI, CT, X-ray, and thermography for the automated diagnosis of breast cancer. We also aim to investigate potential correlations among these different modalities to enhance cancer prediction. In this work, we concentrated on regions of interest (ROIs) from biopsy slides. However, with access to a whole slide scanner, we intend to work with entire cytology slides in future studies. This would

significantly facilitate the collection of a large volume of cytology samples, supporting the development of more robust diagnostic models.

In the future, we aim to further enhance the performance of the automated malignancy identification system. Alongside software improvements, we will explore hardware-based automation to achieve faster and more accurate diagnoses. The work presented in this thesis paves the way for future research in this domain and holds significant societal relevance, with the potential to contribute to a reduction in breast cancer mortality rates.

Thus, as a concluding remark, notable success has been achieved in the identification of malignancy from breast cytology images. In this respect, several computer vision techniques, primarily deep learning-based methods, have made significant contributions toward addressing the research gaps. However, challenges remain, and it would be fascinating for budding researchers to further explore digital pathology images (cytology/histopathology) for breast cancer diagnosis, aiming to reach the next level of success.

Bibliography

[1] M. R. Abbasniya, S. A. Sheikholeslamzadeh, H. Nasiri, and S. Emami, "Classification of breast tumors based on histopathology images using deep features and ensemble of gradient boosting methods," *Computers and Electrical Engineering*, vol. 103, p. 108382, 2022.

[2] B. Abhisheka, S. K. Biswas, B. Purkayastha, D. Das, and A. Escargueil, "Recent trend in medical imaging modalities and their applications in disease diagnosis: a review," *Multimedia Tools and Applications*, vol. 83, no. 14, pp. 43 035–43 070, 2024.

[3] M. A. Agnihotri, K. S. Kothari, and L. P. Naik, "Fat necrosis of the breast masquerading as malignancy diagnosed on fine-needle aspiration cytology," *Journal of Mid-life Health*, vol. 11, no. 1, pp. 49–50, 2020.

[4] F. Ahmad, "Fine needle aspiration cytology fnac: A comprehensive guide to techniques, equipment and clinical applications," *Cytology*, p. 18.

[5] A. M. Al Nemer, "Cytology," PathologyOutlines.com, 2025, accessed: 2025-04-28. [Online]. Available: https://www.pathologyoutlines.com/topic/breastcytology.html

[6] J. Alyami, T. Sadad, A. Rehman, F. Almutairi, T. Saba, S. A. Bahaj, and A. Alkhurim, "Cloud computing-based framework for breast tumor image classification using fusion of alexnet and glcm texture features with ensemble multi-kernel support vector machine (mk-svm)," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, p. 7403302, 2022.

[7] T. Araújo *et al.*, "Classification of breast cancer histology images using convolutional neural networks," *PloS one*, vol. 12, no. 6, p. e0177544, 2017.

[8] M. Arbel, D. J. Sutherland, M. Bińkowski, and A. Gretton, "On gradient regularizers for mmd gans," *Advances in neural information processing systems*, vol. 31, 2018.

[9] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[10] J. Barclay, V. Ernster, K. Kerlikowske, D. Grady, and E. A. Sickles, "Comparison of risk factors for ductal carcinoma in situ and invasive breast cancer," *Journal of the National Cancer Institute*, vol. 89, no. 1, pp. 76–82, 1997.

[11] S. Barratt and R. Sharma, "A note on the inception score," *arXiv preprint arXiv:1801.01973*, 2018.

[12] M. Bińkowski, D. J. Sutherland, M. Arbel, and A. Gretton, "Demystifying mmd gans," *arXiv preprint arXiv:1801.01401*, 2018.

[13] A. Bissoto, F. Perez, E. Valle, and S. Avila, "Skin lesion synthesis with generative adversarial networks," in *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis*. Springer, 2018, pp. 294–302.

[14] V. G. Buddhavarapu *et al.*, "An experimental study on classification of thyroid histopathology images using transfer learning," *Pattern Recognition Letters*, vol. 140, pp. 1–9, 2020.

[15] H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, and M. Wang, "Swin-unet: Unet-like pure transformer for medical image segmentation," in *European conference on computer vision*. Springer, 2022, pp. 205–218.

[16] N. Cao and Y. Liu, "High-noise grayscale image denoising using an improved median filter for the adaptive selection of a threshold," *Applied Sciences*, vol. 14, no. 2, p. 635, 2024.

[17] G. Chen, G. Zhang, Z. Yang, and W. Liu, "Multi-scale patch-gan with edge detection for image inpainting," *Applied Intelligence*, vol. 53, no. 4, pp. 3917–3932, 2023.

[18] X. Chen, X. Zhou, and S. T. Wong, "Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy," *IEEE transactions on biomedical engineering*, vol. 53, no. 4, pp. 762–766, 2006.

[19] K. Das, S. P. K. Karri, A. G. Roy, J. Chatterjee, and D. Sheet, "Classifying histopathology whole-slides using fusion of decisions from deep convolutional network on a collection of random multi-views at multi-magnification," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE, 2017, pp. 1024–1027.

[20] S. Delaloge, S. A. Khan, J. Wesseling, and T. Whelan, "Ductal carcinoma in situ of the breast: finding the balance between overtreatment and undertreatment," *The Lancet*, 2024.

[21] S. Dey, S. Chakraborty, U. G. Roy, and N. Das, "Fuzzy rank-based late fusion technique for cytology image segmentation," in *International Conference on Data, Electronics and Computing*. Springer, 2023, pp. 11–23.

[22] S. Dey, S. Das, S. Ghosh, S. Mitra, S. Chakrabarty, and N. Das, "Syncgan: Using learnable class specific priors to generate synthetic data for improving classifier performance on cytological images," in *Computer Vision, Pattern Recognition, Image Processing, and Graphics: 7th National Conference, NCVPRIPG 2019, Hubballi, India, December 22–24, 2019, Revised Selected Papers 7*. Springer, 2020, pp. 32–42.

[23] Q. Duan, E. D. Angelini, and A. F. Laine, "Real-time segmentation by active geometric functions," *Computer methods and programs in biomedicine*, vol. 98, no. 3, pp. 223–230, 2010.

[24] B. S. Ducatman, *Cytology: Diagnostic principles and clinical correlates*. Elsevier, 2020.

[25] W. Fang, F. Zhang, V. S. Sheng, and Y. Ding, "A method for improving cnn-based image recognition using dcgan." *Computers, Materials & Continua*, vol. 57, no. 1, 2018.

[26] A. M. Filho, M. Laversanne, J. Ferlay, M. Colombet, M. Piñeros, A. Znaor, D. M. Parkin, I. Soerjomataram, and F. Bray, "The globocan 2022 cancer estimates: Data sources, methods, and a snapshot of the cancer burden

worldwide," *International Journal of Cancer*, vol. 156, no. 7, pp. 1336–1346, 2025.

[27] P. Filipczuk, T. Fevens, A. Krzyzak, R. Monczak, A. Krzy, and R. Monczak, "Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies," *IEEE Transactions on Medical Imaging*, vol. 32, no. 12, pp. 2169–2178, 2013.

[28] P. Filipczuk, B. Krawczyk, and M. Woźniak, "Classifier ensemble for an effective cytological image analysis," *Pattern Recognition Letters*, vol. 34, no. 14, pp. 1748–1757, 2013.

[29] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "Gan-based synthetic medical image augmentation for increased cnn performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321–331, 2018.

[30] T. Friedrich and S. Menzel, "Standardization of gram matrix for improved 3d neural style transfer," in *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2019, pp. 1375–1382.

[31] H. Garud, S. P. K. Karri, D. Sheet, J. Chatterjee, M. Mahadevappa, A. K. Ray, A. Ghosh, and A. K. Maity, "High-Magnification Multi-views Based Classification of Breast Fine Needle Aspiration Cytology Cell Samples Using Fusion of Decisions from Deep Convolutional Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017, pp. 828–833.

[32] Y. M. George, H. H. Zayed, M. I. Roushdy, and B. M. Elbagoury, "Remote Computer-Aided Breast Cancer Detection and Diagnosis System Based on Cytological Images," *IEEE Systems Journal*, vol. 8, no. 3, pp. 949–964, 2014.

[33] S. Ghosh, N. Das, and M. Nasipuri, "Reshaping inputs for convolutional neural network: Some common and uncommon methods," *Pattern Recognition*, vol. 93, pp. 79–94, 2019.

[34] D. Gomes Pinto and F. C. Schmitt, "Overcoming pitfalls in breast fine-needle aspiration cytology: a practical review," *Acta Cytologica*, vol. 68, no. 3, pp. 206–218, 2024.

[35] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[36] Q. Guan, Y. Wang, B. Ping, D. Li, J. Du, Y. Qin, H. Lu, X. Wan, and J. Xiang, "Deep convolutional neural network vgg-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: a pilot study," *Journal of Cancer*, vol. 10, no. 20, p. 4876, 2019.

[37] X. Guan, S. Jian, P. Hongda, Z. Zhiguo, and G. Haibin, "An image enhancement method based on gamma correction," in *2009 Second international symposium on computational intelligence and design*, vol. 1. IEEE, 2009, pp. 60–63.

[38] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, and S. Li, "Breast cancer multi-classification from histopathological images with structured deep learning model," *Scientific reports*, vol. 7, no. 1, p. 4172, 2017.

[39] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[40] M. Hrebień, J. Korbicz, and A. Obuchowicz, "Hough transform, (1+1) search strategy and watershed algorithm in segmentation of cytological images," in *Advances in Soft Computing*, M. Kurzynski, E. Puchala, M. Wozniak, and A. Zolnierek, Eds., vol. 45. Springer, 2007, pp. 550–557.

[41] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[42] A. G. Hussien, A. Pop, S. Kumar, F. A. Hashim, and G. Hu, "A novel artificial electric field algorithm for solving global optimization and real-world engineering problems," *Biomimetics*, vol. 9, no. 3, p. 186, 2024.

[43] N. A. M. Isa, E. Subramaniam, M. Y. Mashor, and N. H. Othman, "Fine needle aspiration cytology evaluation for classifying breast cancer using artificial neural network," *American Journal of Applied Sciences*, vol. 4, no. 12, pp. 999–1008, 2007.

[44] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.

[45] S. Issac Niwas, P. Palanisamy, K. Sujathan, E. Bengtsson, S. I. Niwas, P. Palanisamy, K. Sujathan, and E. Bengtsson, "Analysis of nuclei textures of fine needle aspirated cytology images for breast cancer diagnosis using Complex Daubechies wavelets," *Signal Processing*, vol. 93, no. 10, pp. 2828–2837, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168412002216

[46] L. Jeleń, T. Fevens, and A. Krzyżak, "Classification of Breast Cancer Malignancy Using Cytological Images of Fine Needle Aspiration Biopsies," *Int. J. Appl. Math. Comput. Sci.*, vol. 18, no. 1, pp. 75–83, 3 2008.

[47] Y. Jing, Y. Yang, Z. Feng, J. Ye, Y. Yu, and M. Song, "Neural style transfer: A review," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 11, pp. 3365–3385, 2019.

[48] M. Kalita, L. B. Mahanta, A. K. Das, and M. Nath, "A new deep learning model with interface for fine needle aspiration cytology image-based breast cancer detection," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 34, no. 3, pp. 1739–1752, 2024.

[49] A. Kashyap, M. Jain, S. Shukla, and M. Andley, "Study of nuclear morphometry on cytology specimens of benign and malignant breast lesions: A study of 122 cases," *Journal of cytology*, vol. 34, no. 1, pp. 10–15, 2017.

[50] V. Kearney, B. P. Ziemer, A. Perry, T. Wang, J. W. Chan, L. Ma, O. Morin, S. S. Yom, and T. D. Solberg, "Attention-aware discrimination for mr-to-ct image translation using cycle-consistent generative adversarial networks," *Radiology: Artificial Intelligence*, vol. 2, no. 2, p. e190027, 2020.

[51] S. U. Khan, N. Islam, Z. Jan, K. Haseeb, S. I. A. Shah, and M. Hanif, "A machine learning-based approach for the segmentation and classification of malignant cells in breast cytology images using gray level co-occurrence matrix (glcm) and support vector machine (svm)," *Neural Computing and Applications*, pp. 1–8, 2022.

[52] S. Khan, N. Islam, Z. Jan, I. U. Din, and J. J. C. Rodrigues, "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning," *Pattern Recognition Letters*, vol. 125, pp. 1–6, 2019.

[53] J. Kim, A. Harper, V. McCormack, H. Sung, N. Houssami, E. Morgan, M. Mutebi, G. Garvey, I. Soerjomataram, and M. M. Fidler-Benaoudia, "Global patterns and trends in breast cancer incidence and mortality across 185 countries," *Nature Medicine*, pp. 1–9, 2025.

[54] S. K. Kopparapu and M. Satish, "Identifying optimal gaussian filter for gaussian noise removal," in *2011 Third National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics*. IEEE, 2011, pp. 126–129.

[55] M. Kowal, P. Filipczuk, A. Obuchowicz, and J. Korbicz, "Computer-aided diagnosis of breast cancer using gaussian mixture cytological image segmentation," *Journal of medical informatics & Technologies*, vol. 17, 2011.

[56] M. Kowal, M. Skobel, A. Gramacki, and J. Korbicz, "Breast cancer nuclei segmentation and classification based on a deep learning approach," *International Journal of Applied Mathematics and Computer Science*, vol. 31, no. 1, pp. 85–106, 2021.

[57] M. Kowal, M. Żejmo, M. Skobel, J. Korbicz, and R. Monczak, "Cell nuclei segmentation in cytological images using convolutional neural network and seeded watershed algorithm," *Journal of Digital Imaging*, vol. 33, no. 1, pp. 231–242, Feb 2020. [Online]. Available: https://doi.org/10.1007/s10278-019-00200-8

[58] P. Kowal, M.and Filipczuk, A. Obuchowicz, and J. Korbicz, "Computer-aided diagnosis of breast cancer using gaussian mixture cytological image segmentation," *Journal of Medical Informatics & Technologies*, vol. 17, pp. 257–262, 2011.

[59] S. Kuiry, N. Das, A. Das, and M. Nasipuri, "Edc3: Ensemble of deep-classifiers using class-specific copula functions to improve semantic image segmentation," *arXiv preprint arXiv:2003.05710*, 2020.

[60] Y. Kurmi, V. Chaurasia, and N. Ganesh, "Tumor malignancy detection using histopathology imaging," *Journal of Medical Imaging and Radiation Sciences*, vol. 50, no. 4, pp. 514–528, 2019.

[61] M. H. Lafta, A. Afra, I. Patra, A. T. Jalil, M. J. Mohammadi, A. M. Baqir Al-Dhalimy, S. Ziyadullaev, F. Kiani, H. A. Ekrami, and P. Asban, "Toxic effects due to exposure heavy metals and increased health risk assessment (leukemia)," *Reviews on Environmental Health*, vol. 39, no. 2, pp. 351–362, 2024.

[62] J. Lapuyade-Lahorgue, S. Ruan, H. Li, and P. Vera, "Tumor segmentation by fusion of mri images using copula based statistical methods," in *2016 IEEE International Conference on Image Processing (ICIP)*, Sep. 2016, pp. 4136–4139.

[63] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681–4690.

[64] K.-m. Lee and N. Street, "Generalized Hough Transforms with Flexible Templates," in *2000 International Conference on Artificial Intelligence (IC-AI'2000)*, Las Vegas, 6 2000, pp. 1133–1139.

[65] Y. Lee, M. R. Alam, H. Park, K. Yim, K. J. Seo, G. Hwang, D. Kim, Y. Chung, G. Gong, N. H. Cho *et al.*, "Improved diagnostic accuracy of thyroid fine-needle aspiration cytology with artificial intelligence technology," *Thyroid*, vol. 34, no. 6, pp. 723–734, 2024.

[66] F. Leon, S.-A. Floria, and C. Bădică, "Evaluating the effect of voting methods on ensemble-based classification," in *2017 IEEE international conference on INnovations in intelligent Systems and applications (INISTA)*. IEEE, 2017, pp. 1–6.

[67] J. Ma, Y. He, F. Li, L. Han, C. You, and B. Wang, "Segment anything in medical images," *Nature Communications*, vol. 15, no. 1, p. 654, 2024.

[68] J. Malek, A. Sebri, S. Mabrouk, K. Torki, and R. Tourki, "Automated Breast Cancer Diagnosis Based on GVF-Snake Segmentation, Wavelet Features Extraction and Fuzzy Classification," *J. Signal Process. Syst.*, vol. 55, no. 1-3, pp. 49–66, 4 2009.

[69] R. Man, P. Yang, and B. Xu, "Classification of breast cancer histopathological images using discriminative patches screened by generative adversarial networks," *IEEE access*, vol. 8, pp. 155 362–155 377, 2020.

[70] E. McAlpine, P. Michelow, E. Liebenberg, and T. Celik, "Is it real or not? toward artificial intelligence-based realistic synthetic cytology image generation to augment teaching and quality assurance in pathology," *Journal of the American Society of Cytopathology*, vol. 11, no. 3, pp. 123–132, 2022.

[71] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *2018 international interdisciplinary PhD workshop (IIPhDW)*. IEEE, 2018, pp. 117–122.

[72] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014.

[73] S. Mitra, N. Das, S. Dey, S. Chakraborty, M. Nasipuri, and M. K. Naskar, "Cytology image analysis techniques toward automation: systematically revisited," *ACM Computing Surveys (CSUR)*, vol. 54, no. 3, pp. 1–41, 2021.

[74] S. Mitra, S. Dey, N. Das, S. Chakrabarty, M. Nasipuri, and M. K. Naskar, "on superpixel and convolutional neural," *Intelligent Computing Paradigm: Recent Trends*, vol. 784, p. 103, 2019.

[75] T. Mouroutis, S. J. Roberts, and A. A. Bharath, "Robust cell nuclei segmentation using statistical modelling," *Bioimaging*, vol. 6, no. 2, pp. 79–91, 1998.

[76] H. Nguyen, "Fast object detection framework based on mobilenetv2 architecture and enhanced feature pyramid," *J. Theor. Appl. Inf. Technol*, vol. 98, no. 05, pp. 812–824, 2020.

[77] H. Noh, A. E. Ghouch, and T. Bouezmarni, "Copula-based regression estimation and inference," *Journal of the American Statistical Association*, vol. 108, no. 502, pp. 676–688, 2013.

[78] O. Ozdemir, T. G. Allen, S. Choi, T. Wimalajeewa, and P. K. Varshney, "Copula based classifier fusion under statistical dependence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 11, pp. 2740–2748, Nov 2018.

[79] H. S. Park, Y. Chong, Y. Lee, K. Yim, K. J. Seo, G. Hwang, D. Kim, G. Gong, N. H. Cho, C. W. Yoo *et al.*, "Deep learning-based computational cytopathologic diagnosis of metastatic breast carcinoma in pleural fluid," *Cells*, vol. 12, no. 14, p. 1847, 2023.

[80] M. E. Plissiti, P. Dimitrakopoulos, G. Sfikas, C. Nikou, O. Krikoni, and A. Charchanti, "Sipakmed: A new dataset for feature and image based classification of normal and pathological cervical cells in pap smear images," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 3144–3148.

[81] R. Polikar, "Ensemble learning," *Ensemble machine learning: Methods and applications*, pp. 1–34, 2012.

[82] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio, "Transfusion: Understanding transfer learning for medical imaging," in *Advances in neural information processing systems*, 2019, pp. 3347–3357.

[83] N. Rajbongshi, K. Bora, D. C. Nath, A. K. Das, and L. B. Mahanta, "Analysis of morphological features of benign and malignant breast cell extracted from fnac microscopic image using the pearsonian system of curves," *Journal of cytology*, vol. 35, no. 2, pp. 99–104, 2018.

[84] S. R. Ramala Jr, S. Chandak, M. S. Chandak, S. Annareddy, and S. Annareddy Jr, "A comprehensive review of breast fibroadenoma: Correlating clinical and pathological findings," *Cureus*, vol. 15, no. 12, 2023.

[85] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, vol. 28, 2015.

[86] M. I. Rimi, Y. Iliyasu, A. M. Yaro, I. Agyigra, and M. Kabir, "A 6-year retrospective study of fine needle aspiration cytology pattern of otorhinolaryngological cases of patients referred to national ear care centre kaduna (2013-2018)," *International Journal of Research in Medical Sciences*, vol. 11, no. 6, p. 1900, 2023.

[87] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer, 2015, pp. 234–241.

[88] S. Rothe, B. Kudszus, and D. Söffker, "Does classifier fusion improve the overall performance? numerical analysis of data and fusion method characteristics influencing classifier fusion performance," *Entropy*, vol. 21, no. 9, p. 866, 2019.

[89] A. R. Saikia, K. Bora, L. B. Mahanta, and A. K. Das, "Comparative assessment of cnn architectures for classification of breast fnac images," *Tissue and Cell*, vol. 57, pp. 8–14, 2019.

[90] K. Santosh, N. Das, and S. Ghosh, *Deep Learning Models for Medical Imaging*. Academic Press, 2021.

[91] T. K. Santoshi, K. Komaravelli, H. Bathina, G. K. Kumari, P. C. Bharathi, and M. R. Kumar, "Comparative analysis of fnac and histopathology in diagnosing salivary gland lesions at tertiary care center," *Int. J. Trop. Med*, vol. 18, pp. 68–71, 2024.

[92] R. Sanyal, D. Kar, and R. Sarkar, "Carcinoma type classification from high-resolution breast microscopy images using a hybrid ensemble of deep convolutional features and gradient boosting trees classifiers," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2021. [Online]. Available: https://doi.org/10.1109/TCBB.2021.3071022

[93] F. Shahidi, "Breast cancer histopathology image super-resolution using wide-attention gan with improved wasserstein gradient penalty and perceptual loss," *IEEE Access*, vol. 9, pp. 32 795–32 809, 2021.

[94] H.-C. Shin, N. A. Tenenholtz, J. K. Rogers, C. G. Schwarz, M. L. Senjem, J. L. Gunter, K. P. Andriole, and M. Michalski, "Medical image synthesis for data augmentation and anonymization using generative adversarial networks," in *International Workshop on Simulation and Synthesis in Medical Imaging*. Springer, 2018, pp. 1–11.

[95] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, Jul. 2019.

[96] A. Singh and R. S. Ohgami, "Super-resolution digital pathology image processing of bone marrow aspirate and cytology smears and tissue sections," *Journal of Pathology Informatics*, vol. 9, no. 1, p. 48, 2018.

[97] F. A. Spanhol, L. S. Oliveira, P. R. Cavalin, C. Petitjean, and L. Heutte, "Deep features for breast cancer histopathological image classification," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2017, pp. 1868–1873.

[98] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 7, pp. 1455–1462, 2016.

[99] W. N. Street, "Xcyt: a System for Remote Cytological Diagnosis and Prognosis of Breast Cancer," in *Artificial Intelligence Techniques in Breast Cancer Diagnosis and Prognosis*, ser. Series in Machine Perception and Artificial Intelligence. World Scientific Publishing Co., 8 2000, vol. 39, pp. 297–326. [Online]. Available: https://doi.org/10.1142/9789812792488_0008

[100] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[101] H. Tang, H. Liu, D. Xu, P. H. Torr, and N. Sebe, "Attentiongan: Unpaired image-to-image translation using attention-guided generative adversarial networks," *IEEE transactions on neural networks and learning systems*, 2021.

[102] Z. Tao, A. Shi, C. Lu, T. Song, Z. Zhang, and J. Zhao, "Breast cancer: epidemiology and etiology," *Cell biochemistry and biophysics*, vol. 72, pp. 333–338, 2015.

[103] A. Teramoto, A. Yamada, T. Tsukamoto, Y. Kiriyama, E. Sakurai, K. Shiogama, A. Michiba, K. Imaizumi, K. Saito, and H. Fujita, "Mutual stain conversion between giemsa and papanicolaou in cytological images using cycle generative adversarial network," *Heliyon*, vol. 7, no. 2, 2021.

[104] S. Thapa, R. H. Phulware, A. Kumar, and S. Kishore, "Cytomorphological features of complex fibroadenoma breast," *Journal of Cytology*, vol. 40, no. 4, pp. 220–222, 2023.

[105] M. Toğaçar, K. B. Özkurt, B. Ergen, and Z. Cömert, "Breastnet: A novel convolutional neural network model through histopathological images for the diagnosis of breast cancer," *Physica A: Statistical Mechanics and its Applications*, vol. 545, p. 123592, 2020.

[106] F. Tom and D. Sheet, "Simulating patho-realistic ultrasound images using deep generative networks with adversarial learning," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, 2018, pp. 1174–1177.

[107] M. S. Uddin, J. Li *et al.*, "Generative adversarial networks for visible to infrared video conversion," *Recent Advances in Image Restoration with Applications to Real World Problems*, pp. 285–289, 2020.

[108] E. E. Ulku and A. Y. Camurcu, "Computer aided brain tumor detection with histogram equalization and morphological image processing techniques," in *2013 International Conference on Electronics, Computer and Computation (ICECCO)*. IEEE, 2013, pp. 48–51.

[109] S. Ullah and S.-H. Song, "Srresnet performance enhancement using patch inputs and partial convolution-based padding." *Computers, Materials & Continua*, vol. 74, no. 2, 2023.

[110] H. Wang, S. Jiang, and Y. Gao, "Improved object detection algorithm based on faster rcnn," in *Journal of Physics: Conference Series*, vol. 2395, no. 1. IOP Publishing, 2022, p. 012069.

[111] Z. Wang, P. Wang, K. Liu, P. Wang, Y. Fu, C.-T. Lu, C. C. Aggarwal, J. Pei, and Y. Zhou, "A comprehensive survey on data augmentation," *arXiv preprint arXiv:2405.09591*, 2024.

[112] P. Welander, S. Karlsson, and A. Eklund, "Generative adversarial networks for image-to-image translation on multi-contrast mr images-a comparison of cyclegan and unit," *arXiv preprint arXiv:1806.07777*, 2018.

[113] B. Weyn, G. van de Wouwer, A. van Daele, P. Scheunders, D. van Dyck, E. van Marck, and W. Jacob, "Automated breast tumor diagnosis and grading based on wavelet chromatin texture description." *Cytometry*, vol. 33, no. 1, pp. 32–40, 9 1998.

[114] K. P. Win, Y. Kitjaidure, K. Hamamoto, and T. Myo Aung, "Computer-assisted screening for cervical cancer using digital image processing of pap smear images," *Applied Sciences*, vol. 10, no. 5, p. 1800, 2020.

[115] W. H. Wolberg, W. N. Street, and O. L. Mangasarian, "Breast cytology diagnosis with digital image analysis." *Analytical and quantitative cytology and histology / the International Academy of Cytology [and] American Society of Cytology*, vol. 15, no. 6, pp. 396–404, 1993.

[116] ——, "Image analysis and machine learning applied to breast cancer diagnosis and prognosis," *Analytical and Quantitative cytology and histology*, vol. 17, no. 2, pp. 77–87, 1995.

[117] M. S. Wong and W. Y. Yan, "Investigation of diversity and accuracy in ensemble of classifiers using bayesian decision rules," in *2008 International Workshop on Earth Observation and Remote Sensing Applications*. IEEE, 2008, pp. 1–6.

[118] Y. Xue, J. Ye, Q. Zhou, L. R. Long, S. Antani, Z. Xue, C. Cornwell, R. Zaino, K. C. Cheng, and X. Huang, "Selective synthetic augmentation with histogan for improved histopathology image classification," *Medical image analysis*, vol. 67, p. 101816, 2021.

[119] X. Yang, H. Li, and X. Zhou, "Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and Kalman filter in time-lapse microscopy," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, no. 11, pp. 2405–2414, 2006.

[120] C. G. Yedjou, S. S. Tchounwou, R. A. Aló, R. Elhag, B. Mochona, and L. Latinwo, "Application of machine learning algorithms in breast cancer diagnosis and classification," *International journal of science academic research*, vol. 2, no. 1, p. 3081, 2021.

[121] Y. Yu, W. Zhang, and Y. Deng, "Frechet inception distance (fid) for evaluating gans," *China University of Mining Technology Beijing Graduate School: Beijing, China*, 2021.

[122] A. Z. Zambom and R. Dias, "A Review of Kernel Density Estimation with Applications to Econometrics," no. 2000, pp. 1–35, 2012.

[123] M. Zejmo, M. Kowal, J. Korbicz, and R. Monczak, "Classification of breast cancer cytological specimen using convolutional neural network," *Journal of Physics: Conference Series*, vol. 783, no. 012060, p. 012060, 2017. [Online]. Available: http://stacks.iop.org/1742-6596/783/i=1/a=012060?key=crossref.c5d09b2dae030f9a92d444b7461b9765

[124] M. Żejmo, M. Kowal, J. Korbicz, and R. Monczak, "Classification of breast cancer cytological specimen using convolutional neural network," : *Journal of Physics: Conf. Series*, vol. 783, no. 012060, pp. 1–11, 2017.

[125] H. Zerouaoui, A. Idri, F. Z. Nakach, and R. E. Hadri, "Breast fine needle cytological classification using deep hybrid architectures," in *Computational Science and Its Applications–ICCSA 2021: 21st International Conference, Cagliari, Italy, September 13–16, 2021, Proceedings, Part II 21*. Springer, 2021, pp. 186–202.

[126] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2881–2890.

[127] L. Zhao and S. Li, "Object detection algorithm based on improved yolov3," *Electronics*, vol. 9, no. 3, p. 537, 2020.

Nibaran Das 13/5/25

Soumyajyoti Dey 13/5/25

Professor
Computer Sc. & Engg. Department
Jadavpur University
Kolkata - 700