OVARIAN CANCER OUTLIER DETECTION AND SUBTYPE CLASSIFICATION

Project Reference No.: *47S_BE_4678*

College ***:*** *K.L.E. Institute of Technology, Hubballi*
Branch ***:*** *Department of Computer Science and Engineering*
Guide(s) ***:*** *Dr. Girish Saunshi*
Student(S) ***:*** *Ms. Sunidhi Patange*
Mr. Sachetan Heralagi
Mr. Vishal Venkatesh Sortur
Ms. Sejal Pore

Keywords:

Deep Learning,Ovarian Cancer,Whole slide Images of Ovarian Carcinoma.

Introduction:

Ovarian cancer remains one of the most challenging diseases to diagnose and treat effectively. Its high mortality rate is partly attributed to late-stage diagnosis and the heterogeneity of the disease, which manifests in various subtypes with distinct molecular characteristics. Subtyping ovarian cancer is crucial for personalized treatment strategies and prognostic assessment. Ovarian cancer, a formidable adversary in the realm of oncology, poses significant challenges due to its elusive nature and diverse manifestations. While advancements in medical science have enhanced our understanding, its high mortality rates persist, primarily due to late-stage diagnoses and the intricate heterogeneity within the disease.Central to addressing these challenges is the imperative task of subclassifying ovarian cancer.

Objectives:

- Build the dataset of ovarian cancer subtypes from diverse histopathology images.
- Develop a methodology to classify ovarian cancer subtypes namely high grade serous carcinoma, clear-cell ovarian carcinoma, endometrioid, low grade serous and mucinous carcinoma on a diverse dataset of histopathology images.
- Develop deep learning model to detect the cancer subtypes at its early stages.

Methodology

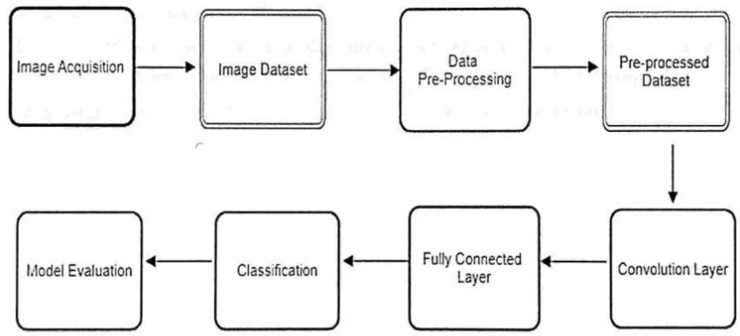

Fig:The proposed achitecture of model

- **Dataset**: Start with a dataset containing samples of data. This could be any type of data such as images,text, numerical values, etc.
- **Data Preprocessing**: This step involves preparing the dataset for the model. It includes tasks like cleaning the data, handling missing values, scaling numerical features, encoding categorical variables,and splitting the data into training and testing sets.
- **Convolutional Layer**: In the context of deep learning, especially for tasks like image recognition, convolutional layers are used to extract features from the input data. These layers apply convolution operations to the input data using learnable filters or kernels! Each filter detects specific patterns or features within the input data.
- **Preprocessed Dataset**: After preprocessing, the dataset is ready to be fed into the model, This includes both the training set, which is used to train the model, and the testing set which is used to evaluate the model's performance.
- **Fully Connected Layer**: Also known as dense layers,fully connected layers are used to perform classification or regression tasks based on the features extracted by the convolutional layers. Each neuron in a fully connected layer is connected to every neuron in the previouslayer, allowing the model tolearn complex patterns in the data.
- **Classification**: Classification is the task of assigning a label or class to each input sample In the context of deep learning, this is typically done using techniques like softmax activation at the output layer, which converts raw predictions into probabilities for each class.
- **Model Evaluation**: Once the model is trained, it needs to be evaluated to assess its performance on unseen data. This involves feeding the testing set into the trained model and calculating metrics such as accuracy, precision, recall, F1 score, etc., depending on the specific task.

In summary, the process involves preparing the dataset, extractíng features using convolutional layers, performing classification using fully connected layers, and evaluating the model's performance

Results and Conclusions:

- The outcome of Ovarian Cancer Subtype Classification and Outlier Detection involves using machine learning to categorize ovarian cancer cases into subtypes and identify unusual cases within the dataset.
- This contributes to improved diagnosis, potential discovery of new subtypes, and early detection of anomalies, enhancing personalized treatment strategies.
- Challenges include interpreting results and ensuring data quality

Scope of future work:

Moving forward, several avenues for future research and development emerge from our study:

- **Enhanced Feature Selection**: Exploring more advanced feature selection techniques to identify the most informative genomic and clinical features could improve the performance of classification models and enhance outlier detection accuracy.
- **Integration of Multi-Omics Data**: Incorporating additional layers of omics data such as proteomics, metabolomics, and epigenomics could provide a more comprehensive understanding of ovarian cancer heterogeneity, enabling more precise classification and outlier detection.
- **Deep Learning Approaches**: Investigating the application of deep learning architectures, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), for ovarian cancer subtype classification and outlier detection could potentially uncover complex patterns within the data that traditional machine learning methods might overlook.
- **Clinical Translation**: Translating our findings into clinical practice by developing user-friendly tools or decision support systems that aid healthcare professionals in accurate subtype classification and outlier identification could significantly impact patient care and treatment planning.
- **Validation Studies**: Conducting independent validation studies on larger and more diverse cohorts to assess the generalizability and robustness of our classification models and outlier detection techniques is crucial for their clinical utility and adoption.
- **Exploration of Therapeutic Implications**: Investigating the therapeutic implications of identified subtypes and outliers, such as exploring targeted therapies or personalized treatment approaches tailored to specific molecular profiles, could pave the way for more effective and personalized ovarian cancer management strategies.

In conclusion, the integration of advanced machine learning techniques with comprehensive omics data holds great promise for advancing our understanding of ovarian cancer heterogeneity and improving patient outcomes. Continued research in this field is essential for realizing the full potential of precision oncology in the fight against ovarian cancer.