Kashf Journal of Multidisciplinary Research
Vol: 02 - Issue 09 (2025)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

UNDERSTANDING URBAN EXPANSION THROUGH MULTI-TEMPORAL SATELLITE DATA ANALYSIS

[1]Salahuddin*, [2]Naeem Aslam, [3]Muhammad Tanveer Meeran, [4]Meiraj Aslam, [5]Muhammad Sajid Maqbool, [6]Bareerah Saeed

[1, 2, 4, 5]Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan.

[3]Faculty of Computer Science and Mathematics, Universiti Malaysia Terengganu, Malaysia.

[6]Department of Computer Science, COMSATS University Islamabad, Vehari Campus.

Corresponding Author: msalahuddin8612@gmail.com

Article Info

Abstract

Satellite imagery represents a vital resource for comprehensively analyzing and monitoring the consequences of rapid urbanization. With the continuous expansion of urban areas, satellite data enables the detection of changes in land use, the spread of urban sprawl, and the development of infrastructure. In the Punjab district of Pakistan, accelerated urban growth has had adverse effects on agricultural land, leading to a decline in agricultural productivity and contributing to a national shortage of food supplies. The utilization of satellite images facilitates the assessment of urbanization's impact on key natural resources, including arable land, forests, wetlands, and river systems. Moreover, satellite-based analysis allows for the annual comparison of land transformation ratios, particularly the conversion of agricultural land into urban settlements. This process aids in identifying land ownership patterns and in understanding the spatial extent and progression of urban expansion. Integrating machine learning classifiers such as Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbor (KNN) with satellite data further enhances the accuracy of applications like land cover classification, change detection, and object recognition. The insights derived from these technologies offer valuable support to policymakers and urban planners, enabling them to develop evidence-based strategies for managing urban growth. Ultimately, such information is instrumental in guiding sustainable urban planning efforts, protecting environmental resources, and prioritizing conservation initiatives in rapidly developing regions.

Keywords:
Urbanization Transformation; Machine Learning; Satellite Imagery; Punjab; Pakistan.

1. INTRODUCTION

In this section, the topic will be explained in clear and precise language, assuming the reader has no prior knowledge of the subject.The introduction aims to establish a friendly rapport with the reader and provide updated and robust information. Images of the Earth captured by satellites and other aircraft-based sensor systems are known as satellite imagery.

These images are valuable for gathering crucial information because they are frequently updated to show how human activities are changing globally. The ongoing process of urbanization keeps many people confined to smaller spaces. As a result of population growth, which has been increasing rapidly over the past few decades due to factors like migration, urbanization, and a rise in fertility rates, urban planning is becoming increasingly important. Urban development can cause a wide range of socioeconomic and environmental issues, both positive and negative.

In recent years, cities have been home to half of the world's population.

This number is expected to surpass one billion by 2025. Rapid, poorly managed urbanization often leads to air, land, and water pollution, as well as inefficient resource usage and the degradation of agricultural land. Changes in land use and land cover have a global impact on the environment. These changes can be caused by either human-made or natural factors. Factors that influence urbanization include population growth, urban expansion, changes in regional environmental characteristics, the loss of agricultural land, and the depletion of natural resources (Zhao et al., 2022). People in countries like Pakistan move from rural to urban areas in search of better facilities, putting pressure on available natural resources. Economic growth makes urbanization inevitable, increasing the demand for housing and the density of urban areas (Busgeeth et al., 2008).

An unprecedented increase in population results in various environmental and social issues, such as inadequate livelihoods, law and order problems, a limited supply of water with declining quality, and climate change (Patra et al., 2018; Murmu et al., 2019).

By 2050, it is anticipated that up to 70% of the world's population will live in urban areas, which raises concern (Zhang, 2016). The migration of thousands of people from rural to urban areas every day (Khan et al., 2024a) in pursuit of better living conditions causes significant pressure on a city's economy and business environment (Lin et al., 2019).

Due to the growth of cities and people, cities expand in both space and physical dimensions (Zlotnik, 2017), leading to increased urban areas, inadequate coverage, and changes in climate and oceans (Khan et al., 2023; Sakieh et al., 2015).

Key problems stemming from unplanned cities include inadequate housing, traffic congestion, basic human services, health issues, unemployment, education (Raza et al., 2024; Al-Khasawneh et al., 2024; Khan et al., 2024b; Farooq et al., 2019), slums, violence, water shortages, power outages, and environmental degradation (Tanguay et al., 2010; Raza et al., 2023; Khan & Asif, 2024; Dai et al., 2024). These factors together significantly impact urban life (Ghalib et al., 2017). Unplanned housing projects resulting from the conversion of agricultural land into urban centers are a major cause of urban expansion in developing countries (Khan & Khan, 2025; Li et al., 2017). Socioeconomic and biophysical factors affect the size and scope of land use and land cover changes in various parts of the world (Farooq et al., 2019). Variations in land use and land cover are some of the most important elements influencing a region's hydrological regime (Nyatuame et al., 2020). Unchecked population growth, rapid urbanization, and spatial changes in agriculture are the main factors influencing land use and cover changes (Shahzad et al., 2024; Hassan et al., 2016; Raza et al., 2023).

In particular, in low- and middle-income countries, unplanned urbanization and rapidly expanding populations often struggle to maintain spatial data and to monitor these changes.

A serious approach towards urban planning and management is necessary, and spatial data on urban sprawl and population growth has become essential, relying heavily on Remote Sensing (RS) and Geographic Information Systems (GIS). Urban sprawl can be measured through traditional methods such as population census data, but these take time. Cost-effective metrics are not efficient for evaluating urban settings. To create integrative plans and mitigation measures, it is necessary to study population growth and urbanization spatially.

Cities have undoubtedly grown due to rapid population growth, economic development, and infrastructure construction activities (Manandhar et al., 2009).

Monitoring and predicting urban growth in terms of settlement planning, transportation, and landscaping is essential for sustainable cities. Currently, we have the ability to obtain necessary data with high geographical, spectral, and temporal resolution.

This study, inspired by Giustarini et al. (2012), aims to analyze the geographic patterns of urbanization and population growth in four major Pakistani cities by assessing the capacity of freely available datasets to understand the dynamics of spatially diverse cityscapes.

The distribution of land use and the spatial growth pattern alongside population density in these four cities will be explored using a variety of sensor data and high-resolution imagery. This research will undoubtedly benefit cost-effective planning in rapidly urbanizing middle-income cities.

Satellite imagery refers to photographs of the Earth or other planets taken from a camera mounted on a satellite orbiting the planet.

These images provide a unique and often highly detailed perspective of the Earth's surface and are used in a wide range of applications, including mapping, monitoring natural disasters, urban planning, and environmental monitoring.

The land serves as the basis for survival, yet humans have continuously altered it for thousands of years, with urbanization accelerating dramatically in recent decades and is expected to continue throughout the century (Hassan, 2016). This rapid transformation of agricultural land has become a central issue in global environmental change and sustainable development. It influences vegetation coverage and can lead to temperature changes and variations in wind speed. It continuously poses a threat to the ecological security of regional land resources [36]. In land cover classification, satellite images are employed to train machine learning models to identify different land cover types, such as forests, urban areas, and agricultural land. The model is trained to recognize the unique characteristics of each land cover type, such as spectral reflectance patterns, texture, and shape.

Main contribution of this work as follow:

The main aims of this study are to analyze the spatial distribution of urbanization and population growth in two districts of Pakistan using satellite imagery, which refers to images captured by satellites orbiting Earth or other planets. Compared to previous years, the rate of urbanization is increasing at a much faster pace today. As a result, it becomes challenging to predict land cover changes in advance.

2. Related work

The machine learning-based approach for classification is considered one of the most effective methods in the field. This is primarily due to the difficulty in obtaining detailed information about the entire area of interest from remote sensing satellite imagery (Lin et al., 2019). Supervised machine learning methods are generally more effective than unsupervised methods because they provide training data necessary for accurate image classification and further analysis (Khan et al., 2023). Selecting the appropriate machine learning classifier presents a significant challenge, largely because of the wide range of available classifiers and the need for both accuracy and efficiency. For instance, Sakieh et al. (2015) found that artificial neural networks (ANNs) were more accurate than decision trees (DTs), although some studies suggest DTs outperform ANNs. The Naive Bayes classifier showed higher accuracy compared to DTs when the number of samples was small, but its performance declined as the number of training samples increased. With smaller training datasets, machine learning classifiers or ANNs outperformed, while support vector machines (SVMs) achieved high accuracy (Raza et al., 2024). Al-Khasawneh et al. (2024) found that both random forests (RF) and SVMs performed nearly equally well in terms of accuracy, with SVMs slightly outperforming RF. Khan et al. (2024b) highlighted that climate change is causing severe droughts globally, and machine learning methods like SVM can be used to predict drought conditions and issue early warning alerts to vulnerable communities. SVM is a popular binary classifier widely used in remote sensing due to its high accuracy with small training data, as these samples are closer to the margin that separates different classes (Farooq & Beg, 2019). The SVM algorithm focuses on finding an optimal hyper-plane that divides training samples into different classes, with samples near the margin serving as support vectors. The use of kernels in SVM plays a crucial role in classification. Among various kernels, the Radial Basis Function (RBF) kernel is preferred due to its user-defined parameters, which help control the influence of training samples on the decision boundary. However, balancing these parameters is essential to prevent overfitting (Tanguay et al., 2010; Raza et al., 2023). Tehrany et al. (2024) conducted trials where SVM was applied in different fields such as flood vulnerability assessment and avalanche studies. Hereditary computing is recognized as a progressive heuristic model in artificial intelligence, with applications in urban planning, environmental exploration, atmospheric studies, and remote sensing. Mueller et al. (2025) indicated that Landsat satellite imagery is beneficial for various applications, such as measuring atmospheric differences and identifying disasters. Landsat data are often used in phenomena-oriented applications. Borra et al. (2026) proposed that satellite image clustering involves collecting image data based on area diversity, with various techniques available for characterizing satellite images to create effective maps. These maps can exhibit different characteristics to achieve optimal results. Several techniques combine guided and unguided approaches to estimate the accuracy of remote sensing data, which is a fundamental aspect of characterization (Nyatuame et al., 2020). Land cover extraction and its changes serve as the initial step in monitoring and assessing urban sustainability. The Maximum Likelihood Classifier (MLC) and post-classification change detection were commonly used in experiments. This methodology highlighted urbanized areas in Hangzhou over the past three years (2001-2003). In a study covering Bangladesh's Dhaka city from 1975 to 2003, the authors examined urban areas. In Kom Ombo, Egypt, the study examined land cover changes from 1998 to 2008 due to agricultural development and urban growth. In Izmir, Turkey, the focus was on urban development. In Samara, Russia, research highlighted land cover changes and population trends over approximately thirty-seven years (1972-2009).

3. Method and materials

The method employed to achieve the study's objectives consists of several steps, as depicted in Figure 1. Raw satellite imagery from the Landsat 8 satellite dataset is used, and various necessary preprocessing steps are applied as outlined below. The United States Geological Survey (USGS) operates under the US Department of the Interior. It is a government laboratory focused on studying Earth's ecosystems. Its aim is to share scientific knowledge with the public to improve disaster preparedness, environmental policy, and resource management. The office also provides information to legislators and community leaders to assist in policy development. Established in 1879, the USGS was originally founded to investigate the geology and geography of the United States. Its findings were intended to help in categorizing public lands for planning. The organization was created to meet the growing needs of the country and has since undertaken a wide range of activities, including environmental mapping, conservation efforts, and surveys of flora and natural water supplies.

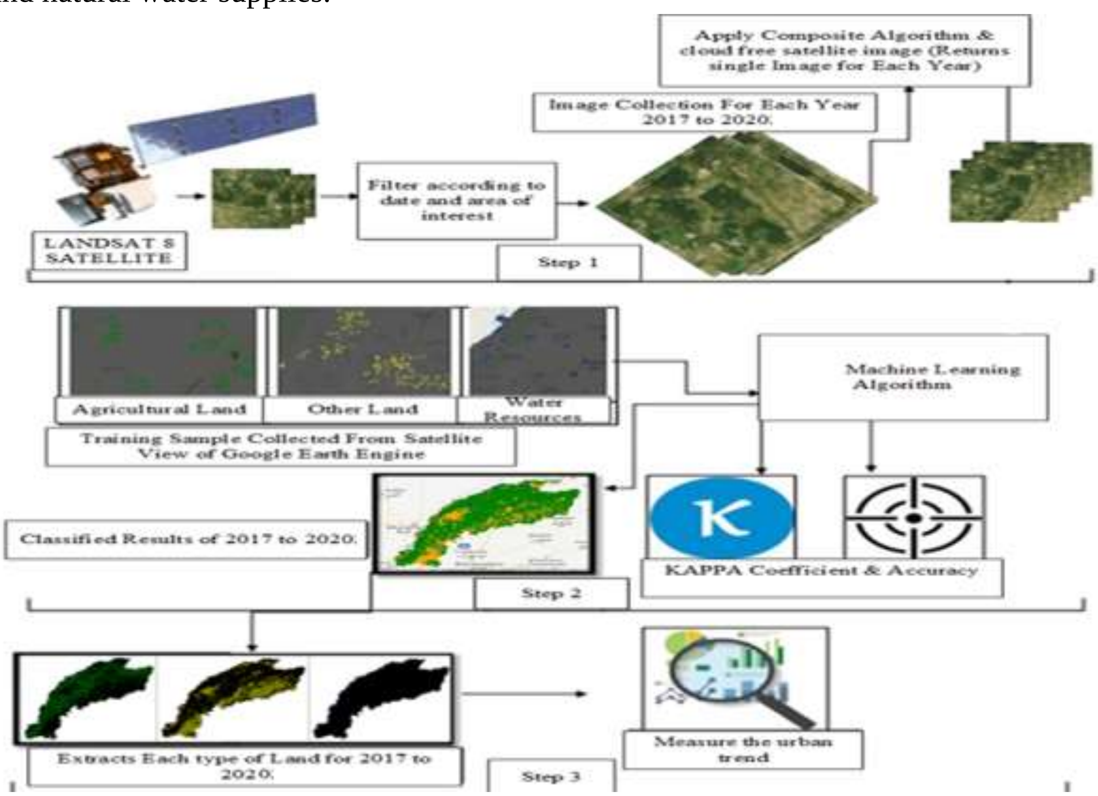

Figure 1 Overview of Proposed Methodology

3.1 Study Area

When discussing urbanization, the term "study area" refers to the specific region under investigation regarding urbanization's patterns, processes, or effects. The extent of the study is determined by the research questions or hypotheses. For this research, the districts of Multan and Khanewal were combined.

Multan is located on the banks of the Chenab River in Punjab, Pakistan.

According to the most recent census (2017), Multan was the seventh most populous city in Pakistan. Multan's coordinates are latitude 30.19838 and longitude 71.4687028000007. Information about Multan is presented below.

Table 1. Multan Attribute Detail

Variable	Details
Latitude DMS	30°11'54.17"N
Longitude DMS	71°28'7.33"E
UTM Easting	737,662.27
UTM Northing	3,343,344.25
UTM Zone	42R
Position from Earth's Center	ENE
Elevation	125.42 Meters (411.482 Feet)
District	Multan
Province	Punjab
Country	Pakistan

- **Khanewal**

Khanewal is a city in the Punjab province of Pakistan and the district capital of the same name. In terms of population, it is only the 36th most populous city in Pakistan. Khanewal, Pakistan may be found at coordinates 30.303934 and 71.9298795, respectively. Information about Khanewal is tabulated below.

Table 2. Khanewal attribute Detail

Variable	Details
Latitude DMS	30°18'14.16"N
Longitude DMS	71°55'47.57"E
UTM Easting	781,775.74
UTM Northing	3,356,101.75
UTM Zone	42R
Position from Earth's Center	ENE
Elevation	136.87 Meters (449.06 Feet)
District	Multan
Province	Punjab
Country	Pakistan

Figure 2 Study Area of this work

3.2 Dataset Preparation

We obtained four annual Landsat TM/OLI images from the United States Geological Survey (https://glovis.usgs.gov/) (Table 1) for both Multan (30°11'54.17N, 71°28'7.33E) and Khanewal (30°18'14.16N, 71°55'47.57E).

The dates of 2017 through 2020 indicate when these images were captured. We will refer to these years from here on as IM2017, IM2018, IM2019, and IM2020. Three or four images were captured every year. The research area's winter and spring crops and other features are at their peak activity from October through December; thus, a four- to five-month interval is necessary between image captures at this time. This helps isolate impermeable surfaces from other forms of land cover. In these images, the cloud cover was around 20% over the research area. We have utilized the QGIS program to create the stack shown in Figure 23 (Bands 2-7) and to capture the three features (Vegetation, Urbanization, Water). After that, we performed radiometric calibrations and atmospheric corrections using the ENVI FLAASH software (Excelis Visual Information Solutions Inc, Boulder, Colorado), as displayed in Figure. We used Google Earth to collect basic geographical data, such as place names (Multan, Khanewal), lakes, green or vegetation area, highways, buildings, airports, and slum regions.

3.3 Preprocessing of Satellite Imagery

Image processing is a method to perform some operations on an image, to get an enhanced image or to extract some useful information from it.

There are some consecutive operations which are performed to preprocess the data consisting of the satellite imagery and make it able for further processing to obtain the described results.

3.3.1 Data Range Filtering

Based on the study area, we have captured raw satellite imagery and filtered it based on a date range (2017-2020).

Specifically, our analysis is based on the years 2017 to 2021, and selects the timeframe from January 1st, 2017, to December 1st, 2020, for each year. To perform this date range filtering operation, we would need to process the satellite imagery data and extract the corresponding dates for each image. Then, you can compare the dates to the specified range to determine if an image falls within the desired timeframe.

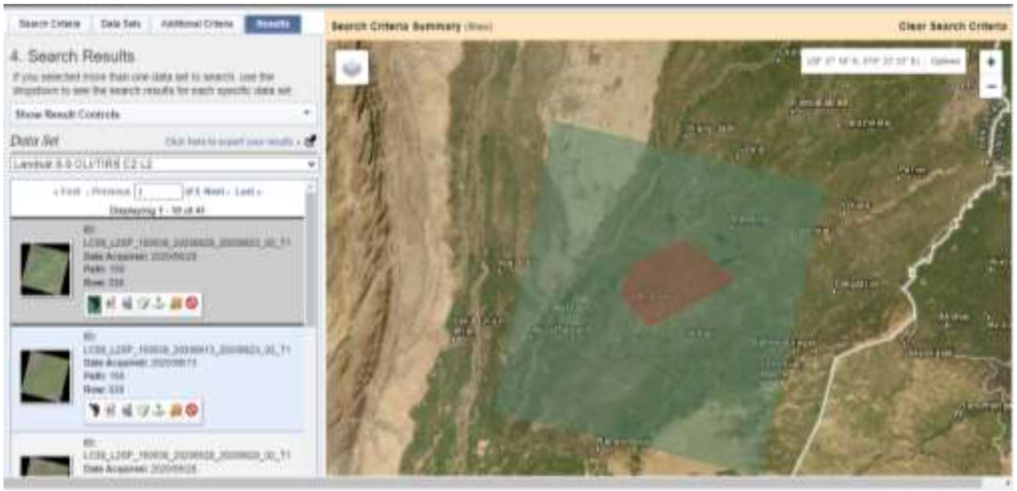

Figure 3. UGIS Sample Filter Data

3.3.2 Area Filter Band

Multan and Khanewal are cities situated in the southern region of Punjab, Pakistan. Here are some additional details about these cities:

- **Multan**

1. Coordinates: 30.1575° N, 71.5249° E

2. Multan is the sixth-largest city in Pakistan and is located on the banks of the Chenab River.

3. Multan is a major cultural and economic center in Punjab and is famous for its handicrafts, pottery, and textile industry.

4. The city has a significant agricultural sector, with major crops including cotton, wheat, sugarcane, and mangoes.

- **Khanewal**

1. Coordinates: 30.2864° N, 71.9320° E

2. Khanewal is a city located in the southern part of Punjab province, near Multan.

3. It is an important transportation hub, as it lies on the main railway line between Karachi and Lahore.

4. Khanewal district is predominantly agricultural, with crops such as wheat, cotton, and sugarcane being grown in the area.

5. The city has a railway junction and is a major stop for trains traveling between the southern and northern parts of Pakistan.

6. Khanewal is also known for its production of quality cotton yarn and textiles.

Both Multan and Khanewal hold historical, cultural, and economic importance in the region, playing a vital role in the development and prosperity of southern Punjab, Pakistan. To filter raw satellite imagery data for a specific administrative area such as the Multan district and its tehsils, you would generally need specialized geographic information system (GIS) software or access to satellite imagery providers or data repositories. These tools can assist in retrieving, filtering, and analyzing satellite imagery based on specific geographical boundaries and parameters.

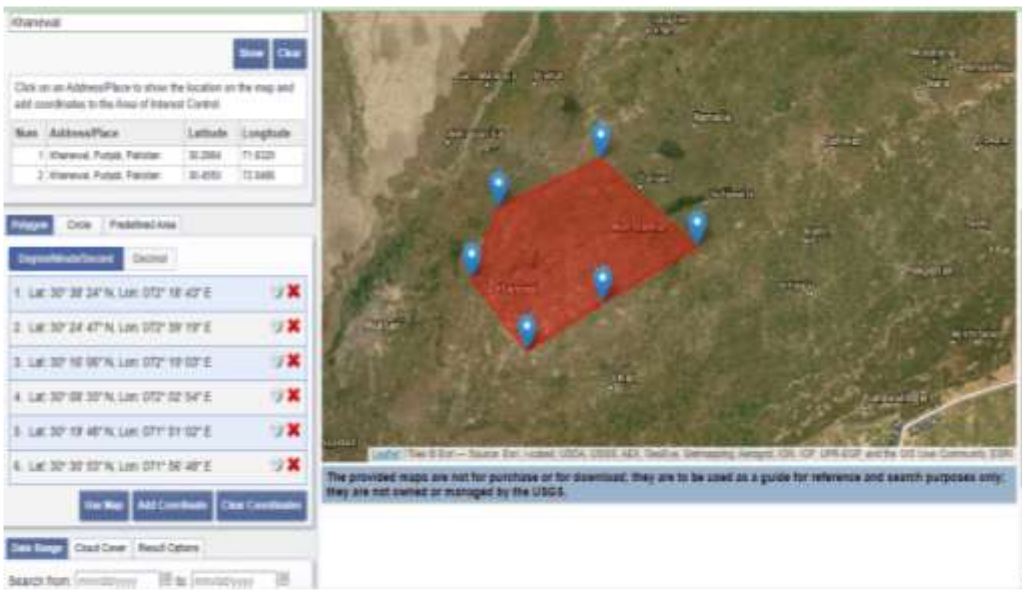

Figure 4 Sample of Study Filter

3.4 Cloud Filtration with Composite Algorithm

The composite algorithm merges spatially overlapping satellite imagery from different years into a single image for each year. This involves combining data from multiple years to create a consolidated image that integrates the best features from each year's imagery. By combining overlapping areas from different years, the composite algorithm aims to produce a single image representing the best available information for each year, leveraging the strengths of each image while minimizing limitations or gaps.

The composite algorithm combines spatially overlapping satellite imagery for the years 2017 to 2020 into a single image per year.

This process uses an aggregation function to select the least cloudy images for further processing.

3.4.1 Collection Training

The training sample points are selected using the ROI tools in ENVI software for the ML Classifier.

The collected training samples are used to train the model, and each sample is labeled based on specific features. Green represents vegetation, dark brown represents agricultural areas, and blue represents water resources. The process involves using the satellite view in ENVI software and applying ML classifiers to label the training samples based on specific features. By utilizing ENVI's satellite view along with ML classifiers and carefully selected training samples, the workflow enables the automated labeling and classification of land cover classes using satellite imagery.

3.4.2 Machine Learning Classifier

In this process, a ML classifier is used to classify raw satellite imagery at the pixel level.

This classification is performed based on labeled training data. The classifier generates classification results for further processing. The results include multiple classes: Vegetation, Water, and Urban. Each image is classified into these three categories. The ML classifier classifies raw satellite imagery based on pixel-level classification, using labeled training data. The classification results are based on multiple classes, and the specific number and types of classes can vary depending on the application and objectives. The ML classifier enables the automated and efficient processing of large-scale satellite imagery, generating accurate and detailed classification results for further analysis and decision-making.

4. Results

Overall accuracy is a measure of the proportion of correctly labeled samples out of the total samples.

It provides a general assessment of the model's performance but may not account for class imbalances or errors in specific classes. Additional evaluation metrics, such as precision, recall, F1 score, or class-specific accuracy, can offer more detailed insights into the model's performance for individual classes. By analyzing the confusion matrix and computing the overall accuracy, the effectiveness of the machine learning model in correctly classifying the samples can be assessed, allowing for informed decisions based on the classification results. The overall accuracy of classification is determined by the confusion matrix, which calculates the proportion of the total sample where the predicted value matches the actual value (measuring the number of correctly labeled samples).

Overall Accuracy = (TP + TN) / (TP + TN + FP + FN)

4.1 Results Evaluation

Machine learning classifiers generate results on a yearly basis, categorizing land into three classes: Urban Land, Agricultural Land, and Water Resources.

These classifiers can be trained to analyze satellite imagery and classify land cover based on features and patterns. Using satellite imagery data, a ML classifier can extract information such as vegetation, built-up areas, and water bodies to distinguish between different land cover types. Urban Land typically refers to areas with buildings, infrastructure, and human settlements. Agricultural Land represents areas used for farming, including cultivated fields and pastures. Water Resources include lakes, rivers, reservoirs, and other water bodies.

The machine learning classifier processes satellite imagery data and assigns each pixel or region to one of the three land cover classes based on its characteristics and identified patterns.

This classification is performed annually to monitor land cover changes over time and analyze trends in urban expansion, agricultural practices, and water resource dynamics. It is important to note that the accuracy of the classification results depends on the quality and resolution of the satellite imagery, the training data provided to the classifier, and the algorithms used. Additionally, manual verification and validation are typically necessary to ensure accuracy and address any misclassifications or errors. Using the classification results from the ML classifier, it is possible to analyze the rate of change in land due to urbanization on a year-by-year basis.

4.1.1 Extraction of Each Kind of Land

To extract each type of land (Urban Land, Agricultural Land, and Water Resources) from the results of the ML classifier for the years 2017 to 2021, ENVI converts the classified pixels into square kilometers.

ENVI is capable of extracting the area of each land cover class for the years 2017 to 2021 based on the similarity ofthe classified pixels and converting them into square kilometers. This information aids in analyzing land cover changes and quantifying the extent of each land cover class over time. From the results of the ML classifier, each type of land (Urban Land, Agricultural Land, and Water Resources) is extracted for the years 2017 to 2021, and the area of each pixel in the classified image is calculated based on similarity, then converted into square kilometers.

4.1.2 Analysis to measure the changing trends

After extracting the area of each land cover class—Urban Land, Agricultural Land, and Water Resources—for the period from 2017 to 2021, we can conduct an analysis to determine the rate of change in land use and understand the changing patterns. This information on urbanization trends can be valuable for various departments and stakeholders. Analyzing the rate of change and the pattern of land cover provides useful insights for different sectors. For example, urban planning departments can use this data to assess the pace and scope of urbanization, agricultural departments can evaluate changes in agricultural land and plan accordingly, and environmental departments can monitor the impact of land use changes on water resources and ecosystems. Once the area of each land cover class has been extracted, apply mathematical methods to calculate the rate of change in land use from 2017 to 2021 and analyze the pattern of change. Understanding urbanization trends is beneficial for multiple departments.

4.2 Study Area (2017-2020)

When considering urbanization, it is essential to clearly define the study area, which refers to the specific geographic region that serves as the focal point for investigating the patterns, processes, or impacts of urbanization. The selection of the study area is typically guided by the research objectives, questions, or hypotheses that need to be addressed. In this research, the study area includes the districts of Multan and

Khanewal. These districts were selected based on the research objectives and their representation of urbanization processes within the study region.

To comprehensively understand urbanization in the study area, it is necessary to analyze and compare data across different time periods.

For this purpose, the ENVI tool, a widely used software for remote sensing and image analysis, was utilized. The tool allows for capturing and visualizing the study area over time, providing valuable insights into the temporal dynamics of urbanization.

The captured figure 5, generated through the ENVI tool, illustrates the study area from 2017 to 2020.

These images likely consist of satellite imagery or other remote sensing data sources processed and analyzed using the ENVI tool's functions. The resulting visual representations enable researchers to observe and quantify changes in urban land cover, infrastructure development, and overall urban growth within the study area over the specified period. By examining these captured images, researchers can identify notable spatial and temporal patterns of urbanization. This may include the expansion of urban areas, the conversion of rural or agricultural land into built-up areas, the emergence of new infrastructure or settlements, and other developments associated with urban growth. These observations can be further analyzed and interpreted to draw conclusions about the processes and impacts of urbanization in the Multan-Khanewal region. Overall, the use of the ENVI tool to capture the study area images from 2017 to 2020 provides researchers with valuable visual data, enabling a detailed assessment of urbanization dynamics over time. This data serves as a foundation for conducting in-depth analysis, modeling, and evaluation of urbanization phenomena in the study area.

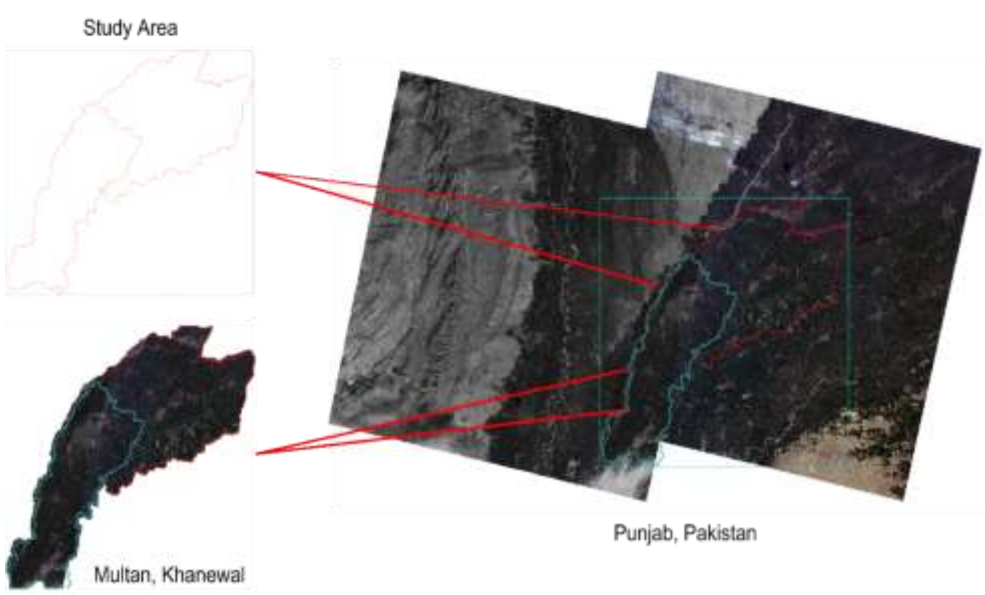

Figure 5 Capture study Area by QGIS Software

4.3 Class Labeling

In this research, the satellite view capability of the ENVI software is used to analyze the study area and investigate urbanization patterns. ENVI provides a robust platform for remote sensing and image analysis, enabling researchers to effectively process and interpret satellite imagery.

To train the ML classifier, the researchers use the ROI tools within ENVI.

These tools allow the selection of specific sample points within the satellite image that represent different land cover classes. Careful selection of these training samples ensures that the ML classifier learns to accurately differentiate between various land cover categories. The selected training samples are used to train the ML model to recognize and classify different land cover classes. The researchers use a set of features extracted from the satellite imagery to label the training samples. These features may include spectral information, such as the intensity of light reflected at different wavelengths, as well as spatial attributes like texture or shape.

Each pixel in the satellite image is assigned an attribute based on the classification from the trained ML model.

In the context of urbanization classification, researchers define specific categories based on their objectives and the characteristics of the study area. For example, they may assign the attribute "green" to represent vegetation, "dark brown" for agricultural areas, and "blue" for water bodies.

By labeling the training samples with these attributes, the ML model learns to associate specific spectral and spatial patterns with each land cover category.

As a result, when applied to the entire satellite image, the ML classifier can accurately classify different land cover classes across the study area.

This approach facilitates a comprehensive analysis of urbanization patterns by distinguishing between vegetation, agricultural areas, and water bodies.

By utilizing the ML classifier trained on the selected samples and the features derived from the satellite imagery, researchers gain insights into the spatial distribution and changes of these land cover classes over time. The integration of ENVI's satellite view, ROI tools, and ML classifier enables a detailed and accurate classification of the study area. By assigning attributes to different land cover categories, the researchers can effectively analyze urbanization patterns and understand the dynamics of vegetation, agriculture, and water resources within the study area. As shown in Figure 6.

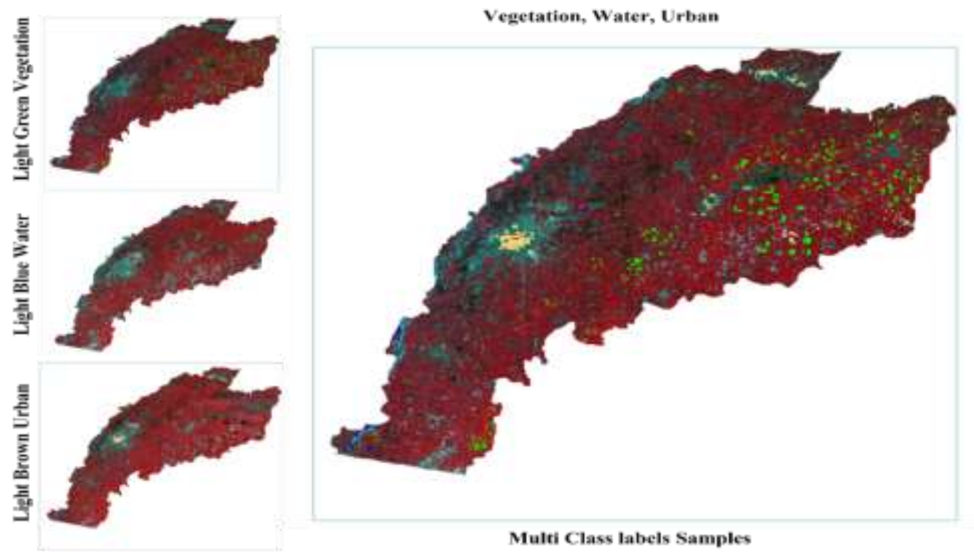

Figure 6 Features Labeling for Classification

4.4 Training KNN Machine Learning Algorithm

In the described process, a machine learning classifier is used to classify raw satellite imagery at the pixel level. This classification assigns each pixel in the satellite image to a specific land cover class, such as Vegetation, Water, or Urban. The classifier's capacity to perform this task is derived from its training on labeled training data sample points.

The labeled training data sample points are carefully selected from the satellite imagery, representing various land cover classes of interest.

These points serve as labeled examples from which the machine learning classifier learns. By analyzing the spectral and spatial characteristics of these samples, the classifier gains an understanding of the unique features associated with each land cover class.

Once the ML classifier is trained, it can be applied to the entire raw satellite imagery.

Each pixel is evaluated based on its spectral values and surrounding context, and the classifier assigns it to one of the predefined land cover classes. This process is referred to as pixel-level classification.

The classification outcomes from this process are valuable for further analysis and processing.

The results provide information on the distribution and spatial extent of different land cover classes within the satellite imagery. For instance, they offer insights into the locations and patterns of vegetation, water bodies, and urban areas. It's important to note that the classification results involve multiple classes, typically more than the three mentioned (Vegetation, Water, Urban).

The ML KNN classifier's ability to classify the raw satellite imagery into multiple classes offers a more detailed understanding of the land cover composition in the study area.

This information can be used for various purposes, such as land use planning, environmental monitoring, or urban growth analysis.

In summary, the process involves using a machine learning classifier to perform pixel-level classification on raw satellite imagery.

Labeled training data sample points allow the classifier to learn the spectral and spatial patterns associated with different land cover classes. The classification results obtained provide valuable insights into the distribution of land cover classes within the imagery, facilitating further analysis and applications across various domains. Figure 7 is showing the training classification of each feature.

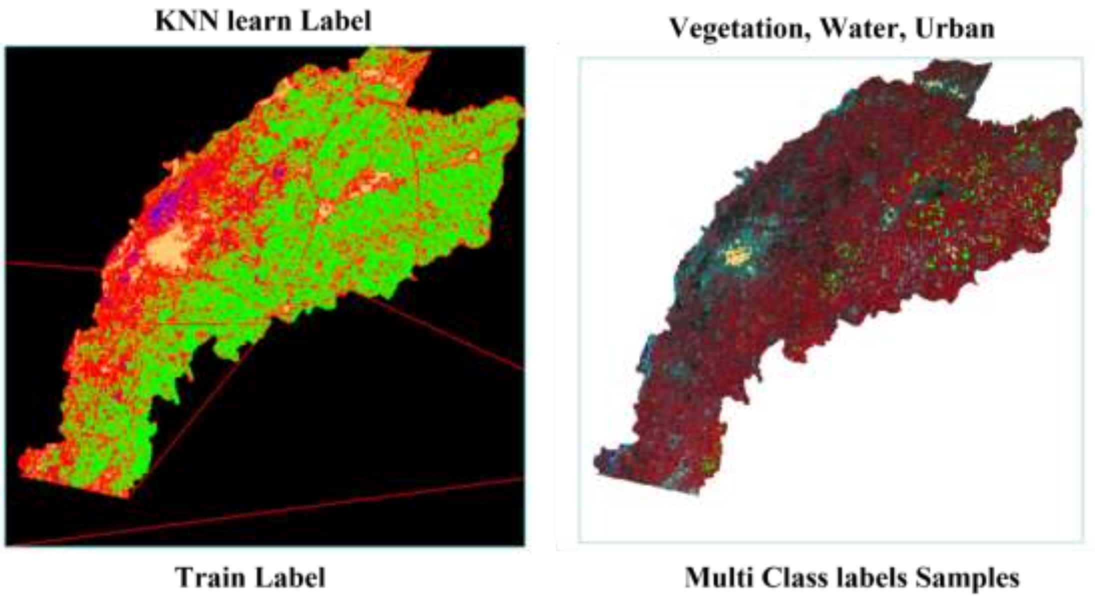

Figure 7 KNN Learn Labels

Table 3 KNN Train Classified Areas and types of land cover and use in Pixels (GT), measured in 2017-2020

Year	Study No	Period	Vegetation Area	Water Area	Urban Area
	S1	01/12 – 03/12	635717	675093	1347875
2017	S2	05/04 – 07/04	1246560	2133146	4999463
	S3	09/09 – 25/12	6378587	264328	1749880
	S1	03/04 – 06/04	6834368	103454	1459301
2018	S2	08/12 – 09/12	6578212	446149	1364238
	S3	12/10 – 12/12	4014892	29529	4346018
	S1	01/17 – 03/17	6405752	834569	1149113
2019	S2	06/27 – 08/27	7142768	45577	1211867

	S3	10/09 – 11/09	6485874	705155	1211207
	S1	01/22 – 03/16	6739925	537049	1123220
2020	S2	08/04 – 09/04	6875440	108490	1400517
	S3	11/13 – 12/25	6113693	745311	1545743

Based on the Vegetation label, Water label, and urban mark assigned during the pixel-level classification process, the resulting classification of the satellite imagery using the KNN algorithm can be compared to the ground truth values. These ground truth values serve as a reference or benchmark against which the accuracy of the KNN classification can be assessed. Table 3, which displays the Pixel Ground Truth values of the KNN Training classification result, presents a comparison between the assigned labels by the KNN algorithm and the actual land cover classes based on ground truth data. This table allows for an evaluation of the accuracy of the KNN classifier in correctly identifying and classifying different land cover features. By analyzing the tables, we can assess the performance of the KNN classifier in terms of accuracy. The accuracy is determined by comparing the assigned labels (Vegetation, Water and Urban) to the ground truth values for each pixel in the satellite imagery. If the assigned label matches the ground truth value, it indicates a correct classification, contributing to the overall accuracy of the classifier. The results presented in the table demonstrate the capability of the KNN algorithm to identify land cover features with acceptable accuracy. The acceptable accuracy achieved by the KNN classifier indicates its effectiveness in distinguishing between Vegetation, Water, and Urban classes within the satellite imagery. This information is valuable for understanding the spatial distribution and extent of these land cover categories and can be utilized in various applications, such as land management, environmental monitoring, or urban planning. In conclusion, the comparison between the Pixel Ground Truth values and the KNN Training classification result provides insights into the accuracy of the KNN classifier in identifying different land cover features. The acceptable accuracy achieved by the classifier demonstrates its effectiveness in distinguishing Vegetation, Water, and Urban classes within the satellite imagery, enabling valuable applications and analysis in various domains. 4.5 Testing Result Table 4 presents the technical details of a study, indicating different periods of observation and the corresponding areas covered by Vegetation, Water, and Urban land cover classes. The values in the table represent the area in square meters and the percentage of the total study area covered by each land cover class during specific periods. For example, Table 4 in Study No "S1" during the period from January 2012 to March 2012, the Vegetation Area covered approximately 1,121,904,000 square meters, which accounted for 14% of the total study area. The Water Area covered approximately 1,919,831,400 square meters, accounting for 9% of the study area. The Urban Area covered approximately 4,499,516,700 square meters, representing 23% of the study area. Similarly, the table provides the respective values for Vegetation, Water, and Urban land cover classes during other study periods (S2 and S3) with their corresponding percentages. These technical details allow researchers and stakeholders to quantitatively analyze and compare the extent of different land cover classes over time, providing insights into the changes in vegetation, water bodies, and urban areas within the study area. Table 4, In addition to the absolute area values, the table includes percentages in parentheses. These percentages represent the proportion of each land cover category relative to the total area of the study area during the corresponding period. For example, in S1, the Vegetation Area covers 26% of the total study area, the Water Area covers 8%, and the Urban Area covers

25%. By presenting the area values and percentages for each land cover category across different study periods, the table allows for the comparison and analysis of changes in land cover over time. Researchers can examine the variations in the extent of vegetation, water bodies, and urban areas, providing insights into the dynamics of land cover patterns and the effects of factors like urbanization, climate, or land management practices. It is important to note that the specific units used for area measurement (e.g., square meters, square kilometers) are not mentioned in the table, so it is assumed that the units are consistent across all values within each column. Additionally, the percentages provide a relative understanding of the land cover composition but do not represent the absolute changes in the area values between study periods. Overall, the table offers a technical representation of the area and proportion of different land cover categories within the study area for each study period, facilitating the analysis and comparison of land cover dynamics over time.

Table 4 Areas and types of land cover and use, measured in 2017-2020

Year	Study No	Period	Vegetation Area	Water Area	Urban Area
	S1	01/12 – 03/12	1121904000.00 (14%)	1919831400.00 (9%)	4499516700 (23%)
2017	S2	05/04 – 07/04	1101904000.00 (10%)	1619831400.00 (7%)	5499516700 (24%)
	S3	09/09 – 25/12	5740728300.00 (25%)	237895200.00 (1%)	5499617770 (24.2%)
	S1	03/04 – 06/04	5920390800.00 (26%)	1819831400.00 (8%)	5590016700 (25%)
2018	S2	08/12 – 09/12	1920390800.00 (16%)	401534100.00 (1%)	5599987700 (25.3%)
	S3	12/10 – 12/12	3613402800.00 (16%)	36576100.00 (0.9%)	5790087722 (27%)
	S1	01/17 – 03/17	1121904000.00 (14%)	1919831400.00 (9%)	5790088800 (27.01%)
2019	S2	06/27 – 08/27	6428491200.00 (28%)	411534700.00 (1%)	5793286700 (27.05%)
	S3	10/09 – 11/09	5837286600.00 (25%)	634639500.00 (2%)	5796244770 (27.08%)
	S1	01/22 – 03/16	6065932500.00 (27%)	483344100.00 (2%)	5796244770 (27.08%)
2020	S2	08/04 – 09/04	6187896000.00 (27%)	976418000.00 (1%)	5799244770 (28%)
	S3	11/13 – 12/25	5502323700.00 (25%)	670779900.00 (3%)	5997777770 (30%)

5. Discussion

This study offers detailed technical data for three distinct study periods (S1, S2, and S3) spanning the years 2017 to 2020. The data is centered around three land cover categories: Vegetation Area, Water Area, and Urban Area. Each table corresponds to a specific study period, with each column representing one of the land cover categories.

For each study period, the table presents numerical values that represent the area covered by each land cover category within the study area.

These area values are measured in appropriate units, such as square meters, and reflect the extent of each land cover category during the respective study period. This information enables a quantitative analysis of the changes and patterns in land cover over time within the study area.

The tables provide valuable insights into the dynamics of land cover classes by presenting the area values for each category across different study periods.

This allows an examination of the shifts and changes in land cover categories over the specified timeframe.

The cross-tabulation matrices, often referred to as the Pixel Ground Truth, further illustrate the nature of changes in the land cover classes.

Specifically, these matrices depict the transitions or shifts between different land cover categories. For instance, the matrices show that within the agriculture class, there was an increase from 10% to 25% between 2017 and 2020. Out of this increase, approximately 25% to 27% originated from a conversion of Vegetation, while the remainder came from changes in Settlement areas.On the other hand, the Water class maintained a relatively small portion throughout each year from 2017 to 2020. Meanwhile, the other land cover classes, such as Urban and potentially other categories not specified, experienced overall increases in their areas each year.

By examining the cross-tabulation matrices, we can gain a deeper understanding of the changes in land cover classes over time and the interactions between different categories. These matrices provide valuable information on the transitions and dynamics within the land cover composition of the study area, shedding light on the processes driving the changes observed. Overall, this study, along with the cross-tabulation matrices, enables the analysis and interpretation of the patterns and changes in land cover over the study period, facilitating a comprehensive understanding of the dynamics and transformations within the study area's land cover composition.

6. Conclusion

Based on the data collected using GIS, QGIS, and ENVI software, we conclude that land cover and land use practices in the study area have undergone significant changes over the past two decades, aligning with the study's research objectives. The transformation in land use within the watershed region is evident, with a smaller proportion of land designated for vegetation and water (3%) and a larger portion used for agriculture (25%) and urban development (30%). The most noticeable consequences of this expansion include deforestation and water scarcity, particularly impacting the Vegetation and Water classes. By 2020, these shifts in land cover and land use had negatively affected water quality and accessibility, which could become a limiting factor in the future. Vegetation cover in watershed areas is declining, with urbanization and agricultural practices likely contributing to this trend. Mapping and analyzing land use change through Remote Sensing and GIS can provide valuable insights. If not properly managed, this vital water resource may be lost, impairing its role in agricultural production and the social and economic development of the area. This study concludes with recommendations for the sustainable management and conservation of the depleting forest, water, and soil resources in the watershed.

Reference

[1] Khan, S.U.R., Asif, S., Bilal, O. et al. Lead-cnn: lightweight enhanced dimension reduction convolutional neural network for brain tumor classification. Int. J. Mach. Learn. & Cyber. (2025). https://doi.org/10.1007/s13042-025-02637-6.

[2] Hekmat, A., et al., Brain tumor diagnosis redefined: Leveraging image fusion for MRI enhancement classification. Biomedical Signal Processing and Control, 2025. 109: p. 108040.

[3] Khan, Z., Hossain, M. Z., Mayumu, N., Yasmin, F., & Aziz, Y. (2024, November). Boosting the Prediction of Brain Tumor Using Two Stage BiGait Architecture. In 2024 International Conference on Digital Image Computing: Techniques and Applications (DICTA) (pp. 411-418). IEEE.

[4] Khan, S. U. R., Raza, A., Shahzad, I., & Ali, G. (2024). Enhancing concrete and pavement crack prediction through hierarchical feature integration with VGG16 and triple classifier ensemble. In 2024 Horizons of Information Technology and Engineering (HITE)(pp. 1-6). IEEE https://doi. org/10.1109/HITE63532.

[5] Khan, S.U.R., Zhao, M. & Li, Y. Detection of MRI brain tumor using residual skip block based modified MobileNet model. Cluster Comput 28, 248 (2025). https://doi.org/10.1007/s10586-024-04940-3

[6] Al-Khasawneh, M. A., Raza, A., Khan, S. U. R., & Khan, Z. (2024). Stock Market Trend Prediction Using Deep Learning Approach. Computational Economics, 1-32.

[7] Khan, U. S., Ishfaque, M., Khan, S. U. R., Xu, F., Chen, L., & Lei, Y. (2024). Comparative analysis of twelve transfer learning models for the prediction and crack detection in concrete dams, based on borehole images. Frontiers of Structural and Civil Engineering, 1-17.

[8] Khan, M. A., Khan, S. U. R., Rehman, H. U., Aladhadh, S., & Lin, D. (2025). Robust InceptionV3 with Novel EYENET Weights for Di-EYENET Ocular Surface Imaging Dataset: Integrating Chain Foraging and Cyclone Aging Techniques. International Journal of Computational Intelligence Systems, 18(1), 1-26.

[9] Khan, S. U. R., Asif, S., Zhao, M., Zou, W., Li, Y., & Xiao, C. (2026). ShallowMRI: A novel lightweight CNN with novel attention mechanism for Multi brain tumor classification in MRI images. Biomedical Signal Processing and Control, 111, 108425.

[10] Khan, S. U. R., Asif, S., Zhao, M., Zou, W., Li, Y., & Li, X. (2025). Optimized deep learning model for comprehensive medical image analysis across multiple modalities. Neurocomputing, 619, 129182.

[11] Khan, S. U. R., Asif, S., Zhao, M., Zou, W., & Li, Y. (2025). Optimize brain tumor multiclass classification with manta ray foraging and improved residual block techniques. Multimedia Systems, 31(1), 1-27.

[12] Khan, Z., Khan, S. U. R., Bilal, O., Raza, A., & Ali, G. (2025, February). Optimizing Cervical Lesion Detection Using Deep Learning with Particle Swarm Optimization. In 2025 6th International Conference on Advancements in Computational Sciences (ICACS) (pp. 1-7). IEEE.

[13] Bilal, O., Hekmat, A., Shahzad, I. et al. Boosting Machine Learning Accuracy for Cardiac Disease Prediction: The Role of Advanced Feature Engineering and Model Optimization. Rev Socionetwork Strat (2025). https://doi.org/10.1007/s12626-025-00190-w

[1**4]** Raza, Asif, Inzamam Shahzad, Muhammad Salahuddin, and Sadia Latif. "Satellite Imagery Employed to Analyze the Extent of Urban Land Transformation in The Punjab District of Pakistan." Journal of Palestine Ahliya University for Research and Studies 4, no. 2 (2025): 17-36.

[15] Asif Raza, Inzamam Shahzad, Ghazanfar Ali, and Muhammad Hanif Soomro. "Use Transfer Learning VGG16, Inception, and Reset50 to Classify IoT Challenge in Security Domain via Dataset Bench Mark." Journal of Innovative Computing and Emerging Technologies 5, no. 1 (2025).

[16] Waqas, M., Bandyopadhyay, R., Showkat Ian, E., Muneer, A., Zafar, A., Alvarez, F. R., ... & Wu, J. (2025). The Next Layer: Augmenting Foundation Models with Structure-Preserving and Attention-Guided Learning for Local Patches to Global Context Awareness in Computational Pathology. ArXiv, arXiv-2508.

[17] Muneer, Amjad, Muhammad Waqas, Maliazurina B. Saad, Eman Showkatian, Rukhmini Bandyopadhyay, Hui Xu, Wentao Li et al. "From Classical Machine Learning to Emerging Foundation Models: Review on Multimodal Data Integration for Cancer Research." arXiv preprint arXiv:2507.09028 (2025).

[18] Khan, S.U.R., Raza, A., Shahzad, I., Khan, S. (2025). Subcellular Structures Classification in Fluorescence Microscopic Images. In: Arif, M., Jaffar, A., Geman, O. (eds) Computing and Emerging Technologies. ICCET 2023. Communications in Computer and Information Science, vol 2056. Springer, Cham. https://doi.org/10.1007/978-3-031-77620-5_20

[19] Hekmat, A., Zuping, Z., Bilal, O., & Khan, S. U. R. (2025). Differential evolution-driven optimized ensemble network for brain tumor detection. International Journal of Machine Learning and Cybernetics, 1-26.

[20] Khan, S. U. R. (2025). Multi-level feature fusion network for kidney disease detection. Computers in Biology and Medicine, 191, 110214.

[21] Khan, S. U. R., Asif, S., & Bilal, O. (2025). Ensemble Architecture of Vision Transformer and CNNs for Breast Cancer Tumor Detection from Mammograms. International Journal of Imaging Systems and Technology, 35(3), e70090.

[22] Raza, A., Salahuddin, & Inzamam Shahzad. (2024). Residual Learning Model-Based Classification of COVID-19 Using Chest Radiographs. Spectrum of Engineering Sciences, 2(3), 367–396.

[23] Raza, A., Soomro, M. H., Shahzad, I., & Batool, S. (2024). Abstractive Text Summarization for Urdu Language. Journal of Computing & Biomedical Informatics, 7(02).

[24] Khan, S. U. R., & Khan, Z. (2025). Detection of Abnormal Cardiac Rhythms Using Feature Fusion Technique with Heart Sound Spectrograms. Journal of Bionic Engineering, 1-20.

[25] Khan, M.A., Khan, S.U.R. & Lin, D. Shortening surgical time in high myopia treatment: a randomized controlled trial comparing non-OVD and OVD techniques in ICL implantation. BMC Ophthalmol 25, 303 (2025). https://doi.org/10.1186/s12886-025-04135-3

[26] Khan, U. S., & Khan, S. U. R. (2025). Ethics by Design: A Lifecycle Framework for Trustworthy AI in Medical Imaging from Transparent Data Governance to Clinically Validated Deployment. arXiv preprint arXiv:2507.04249.

[27] Khan, S. U. R., Asif, S., Zhao, M., Zou, W., Li, Y., & Xiao, C. (2026). ShallowMRI: A novel lightweight CNN with novel attention mechanism for Multi brain tumor classification in MRI images. Biomedical Signal Processing and Control, 111, 108425.

[28] Bilal, O., Hekmat, A., & Khan, S. U. R. (2025). Automated cervical cancer cell diagnosis via grid search-optimized multi-CNN ensemble networks. Network Modeling Analysis in Health Informatics and Bioinformatics, 14(1), 67.

[29] Khan, M. A., Khan, S. U. R., Rehman, H. U., Aladhadh, S., & Lin, D. (2025). Robust InceptionV3 with Novel EYENET Weights for Di-EYENET Ocular Surface Imaging Dataset: Integrating Chain Foraging and Cyclone Aging Techniques. International Journal of Computational Intelligence Systems, 18(1), 204.

[30] Maqsood, H., & Khan, S. U. R. (2025). MeD-3D: A Multimodal Deep Learning Framework for Precise Recurrence Prediction in Clear Cell Renal Cell Carcinoma (ccRCC). arXiv preprint arXiv:2507.07839.

[31] Khan, S. R., Asif Raza, Inzamam Shahzad, & Hafiz Muhammad Ijaz. (2024). Deep transfer CNNs models performance evaluation using unbalanced histopathological breast cancer dataset. Lahore Garrison University Research Journal of Computer Science and Information Technology, 8(1).

[32] O. Bilal, Asif Raza, S. ur R. Khan, and Ghazanfar Ali, “A Contemporary Secure Microservices Discovery Architecture with Service Tags for Smart City Infrastructures ”, VFAST trans. softw. eng., vol. 12, no. 1, pp. 79–92, Mar. 2024

[33] S.ur R. Khan, Asif. Raza, Muhammad Tanveer Meeran, and U. Bilhaj, “Enhancing Breast Cancer Detection through Thermal Imaging and Customized 2D CNN Classifiers”, VFAST trans. softw. eng., vol. 11, no. 4, pp. 80–92, Dec. 2023.

[34] Latif, Sadia, Sami Ullah, Aafia Latif, Ghazanfar Ali, Muhammad Hassnain Azhar, and Salman Ali. "PREDICTIVE MODELING OF CARDIOVASCULAR DISEASE USING MACHINE LEARNING APPROACH." Kashf Journal of Multidisciplinary Research 2, no. 02 (2025): 207-232.

[35] Latif, Aafia, Sadia Latif, and Furqan Jamil. "UTILIZING BIG DATA ANALYTICS TO OPTIMIZE INFORMATION COMMUNICATION TECHNOLOGY STRATEGIES." Kashf Journal of Multidisciplinary Research 1, no. 12 (2024): 122-140.

[36] Latif, Sadia, Azhar Mehboob, Muhammad Ramzan, and Muhammad Ans Khalid. "DETECTION OF HCV LIVER FIBROSIS APPLYING MACHINE LEARNING TECHNIQUE." Kashf Journal of Multidisciplinary Research 1, no. 12 (2024): 11-44.

[37] Irtaza, Ghulam, Sadia Latif, Rana Muhammad Nadeem, Nasir Hussain, and Hafiz Muhammad Ijaz. "Real-time Satellite Image Classification Using Convolutional Neural Networks for Earth Observation and Video Feed Analysis." Kashf Journal of Multidisciplinary Research 2, no. 03 (2025): 204-216.

[38] HUSSAIN, S., Raza, A., MEERAN, M. T., IJAZ, H. M., & JAMALI, S. (2020). Domain Ontology Based Similarity and Analysis in Higher Education. IEEEP New Horizons Journal, 102(1), 11-16.

[39] Mahmood, F., Abbas, K., Raza, A., Khan,M.A., & Khan, P.W. (2019). Three Dimensional Agricultural Land Modeling using Unmanned Aerial System (UAS). International Journal of Advanced Computer Science and Applications (IJACSA) [p-ISSN : 2158-107X, e-ISSN : 2156-5570], 10(1).

[40] Raza, A., & Meeran, M. T. (2019). Routine of Encryption in Cognitive Radio Network. Mehran University Research Journal of Engineering and Technology [p-ISSN: 0254-7821, e-ISSN: 2413-7219], 38(3), 609-618.

[41] Meeran, M. T., Raza, A., & Din, M. (2018). Advancement in GSM Network to Access Cloud Services. Pakistan Journal of Engineering, Technology & Science [ISSN: 2224-2333], 7(1).

[42] Khan, S. U. R., Asim, M. N., Vollmer, S., & Dengel, A. (2025). FOLC-Net: A Federated-Optimized Lightweight Architecture for Enhanced MRI Disease Diagnosis across Axial, Coronal, and Sagittal Views. arXiv preprint arXiv:2507.06763.

[43] Khan, S. U. R., Asim, M. N., Vollmer, S., & Dengel, A. (2025). Flora Syntropy-Net: Scalable Deep Learning with Novel Flora Syntropy Archive for Large-Scale Plant Disease Diagnosis. arXiv preprint arXiv:2508.17653.