**Supplementary table 1: The characteristics of the dataset**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Dataset** | **Origin (YOI)** | **Patients Number** | **Finding Types** | **Annotation** | **Bi-RADS** | **Breast density** | **Age** |
| **DDSM** | USA (1996) | 2620 | Masses & Calcifications | ROI | Yes | Yes | Yes |
| **CBIS-DDSM** | USA (2017) | 1566 | Masses & Calcifications | ROI | Yes | Yes | No |
| **Mini-DDSM** | USA (2020) | 2620 | Masses & Calcifications | No | No | Yes | Yes |
| **INbreast** | Portugal (2012) | 115 | Calcification, masses. Architectural distortion, asymmetry, and undefined | ROI | Yes | Yes | No |
| **MIAS** | UK (1994) | 161 | Calcification, masses. Architectural distortion, asymmetry, and normal | A circle enclosing the abnormality. | No | Yes | No |
| **CMMD** | China (2021) | 1775 | Calcifications and masses | No | No | No | Yes |
| YOI: Year of Introduction | | | | | | | |

**Supplementary 1** presents the characteristics of the dataset, including its origin, year of introduction, number of patients and images, and the reported findings. It also indicates whether age, BI-RADS assessment, and breast density reporting are included in the dataset, denoted by "Yes" or "No."

**Supplementary table 2:** The distribution of breast lesions

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Dataset** | **Mass** | **Calcifications** | **Both** | **Benign** | **Malignant** | **Nipple Retraction** | **Skin retraction** | **Skin thickening** |
| **CBIS-DDSM** | 1696 | 1872 | NR | 2111 | 1457 | NR | NR | NR |
| **Mini-DDSM** | NR | NR | NR | 3360 | 3596 | NR | NR | NR |
| **INbreast** | 108 | 308 | NR | NR | NR | NR | NR | NR |
| **MIAS** | 59 | 30 | NR | 69 | 54 | NR | NR | NR |
| **CMMD1** | 726 | 158 | 223 | 544 | 563 | NR | NR | NR |
| **CMMD2** | 417 | 98 | 234 | 0 | 749 | NR | NR | NR |

**Supplementary 2** displays the distribution of breast lesions categorized as mass vs. Calcifications and Benign vs. malignant across clinical manifestations (nipple retraction, skin retraction, and skin thickening) for each database.

**Supplementary table 3:The number of images in various views (CC vs. MLO) and distinguishes between the right and left views for each dataset.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Database** | Breast View | | | | Breast Density | | | |
| Right CC | Right MLO | Left CC | Left MLO | A | B | C | D |
| **CBIS-DDSM\*** | 810 | 939 | 862 | 957 | 536 | 1373 | 1103 | 554 |
| **Mini-DDSM** | 1952 | 1952 | 1952 | 1952 | 1076 | 3020 | 2320 | 1388 |
| **INbreast** | 99 | 101 | 104 | 105 | 136 | 146 | 99 | 28 |
| **MIAS** | NA | 161 | NA | 161 | NR | NR | NR | NR |
| **CMMD1** | 514 | 514 | 593 | 593 | NR | NR | NR | NR |
| **CMMD2\*\*** | 370 | 370 | 395 | 395 | NR | NR | NR | NR |
| MLO: mediolateral oblique, CC: craniocaudal  \*Distribution of the images in the training and test sets not the entire database.  \*\*In the dataset description, the total number of images in CMMD2 is mentioned as 1498. However, in the Excel file, the number reported is 1530. | | | | | | | | |

**Supplementary 3** shows the number of images in various views (CC vs. MLO) and distinguishes between the right and left views for each dataset. In addition, it provides breast density classification, ranging from A to D according to the ACR classification system, for each dataset. This classification categorizes the breasts as almost entirely fatty (category A), scattered areas of fibroglandular density (category B), heterogeneously dense (category C), or extremely dense (category D) 30.

**Supplementary table 4:**Tthe distribution of BI-RADS assessments for each database

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Dataset** | **BI-RADS** | | | | | | | **Total** |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| CBIS-DDSM | 238 | 3 | 644 | 477 | 1633 | 577 | NR | 3572 |
| Mini-DDSM | NR | NR | NR | NR | NR | NR | NR | NR |
| INbreast | NR | 67 | 244 | 25 | 56 | 95 | 14 | 501 |

**Supplementary 4** presents the distribution of BI-RADS assessments for each database. It ranges from a BI-RADS score of 0 to 6, where 0 indicates incomplete assessment, 1 indicates negative clinical and radiological findings, 2 indicates benign findings, 3 is probably benign findings, 4 is suspicious malignancy, 5 is highly positive for malignancy, and 6 is biopsy-proven malignancy.