Spren Camera Based Body Composition Validation Study

May 23, 2024

Dr. Kamiar Kordari, Kéanu Spies (Spren / Elite HRV Inc.);
Professor Steven Heymsfield, Cassidy McCarthy (Pennington Biomedical Research Center, Louisiana State University)

Abstract

This paper is a study to evaluate the performance of Spren's body composition app, which uses state-of-the-art computer vision and deep learning to estimate body fat percentage from smartphone camera images. Body fat measurement is a critical indicator of overall health and fitness; however, commercially available tools such as Dual-energy X-ray Absorptiometry (DXA) often have challenges related to accessibility, cost, and invasiveness. This study highlights the efficacy of Spren's non-invasive approach, emphasizing its potential to democratize body composition assessment. The app's algorithm has been validated through data collection from a heterogeneous participant cohort, demonstrating a mean absolute error (MAE) of 2.6 and a correlation of 0.95 with DXA measurements, indicating high concordance with gold-standard methods. The findings suggest that Spren's app offers a reliable, safe, and convenient alternative for body composition analysis, with significant implications for public health monitoring.

Data Collection and Analysis

Participant Pool

To validate the accuracy and reliability of Spren's body composition app, data was collected from a diverse participant pool at Louisiana State University's Pennington Biomedical Research Center. A total of 84 subjects were included, ensuring a broad representation across different ages, genders, body types, and ethnic backgrounds.

Participant Demographics

- **Total number of subjects:**
  - 84
- **Gender composition:**

    - Female: 46 (54.8%)
    - Male: 38 (45.2 %)
- **Ethnic and racial composition:**
    - White or Caucasian: 65 (77.4%)
    - Black or African American: 15 (17.9%)
    - South Asian: 2 (2.4%)
    - Latin American Hispanic: 2 (2.4%)
- **Mean age:** 50.5 ± 16.0 years (range: 23.0–78.0 years)
- **BMI:** 29.3 ± 7.0 kg/m² (range: 17.4–48.5 kg/m²)
- **DXA-measured body fat percentage (%BF):**
    - Women: 40.4 ± 8.00%
    - Men: 29.3 ± 8.38%

Each subject underwent body composition scanning using three different DXA machines: the Hologic Horizon, Hologic Discovery, and GE Lunar iDXA. These devices were chosen for their high accuracy and widespread use in clinical and research settings. By using multiple DXA machines, we aimed to compare the accuracy of Spren by reflecting the variations that might occur across different hardware.

In addition to the DXA scans, each subject was scanned twice using the Spren app to estimate their body fat percentage. This dual-scanning protocol was implemented to assess the repeatability of the Spren app's measurements and to account for any potential variability in image capture. The Spren app scans were conducted under controlled conditions to ensure consistency in lighting, positioning, and other factors that could influence the accuracy of the image analysis.

Results and Validation

DEXA Machine Comparison Analysis

The following table provides summary statistics for body fat percentage measurements obtained using three different DXA machines: Hologic Horizon, GE Lunar iDXA, and Hologic Discovery.

Mean Absolute Error (MAE)
The Mean Absolute Error (MAE) between body fat percentage measurements from different DXA machines is as follows:

- MAE between 'Hologic Horizon' and 'GE Lunar iDXA': 2.68
- MAE between 'Hologic Horizon' and 'Hologic Discovery': 1.95
- MAE between 'GE Lunar iDXA' and 'Hologic Discovery': 2.35

Correlation Matrix
The correlation matrix indicates the Pearson correlation between the body fat percentage measurements from the three DXA machines.

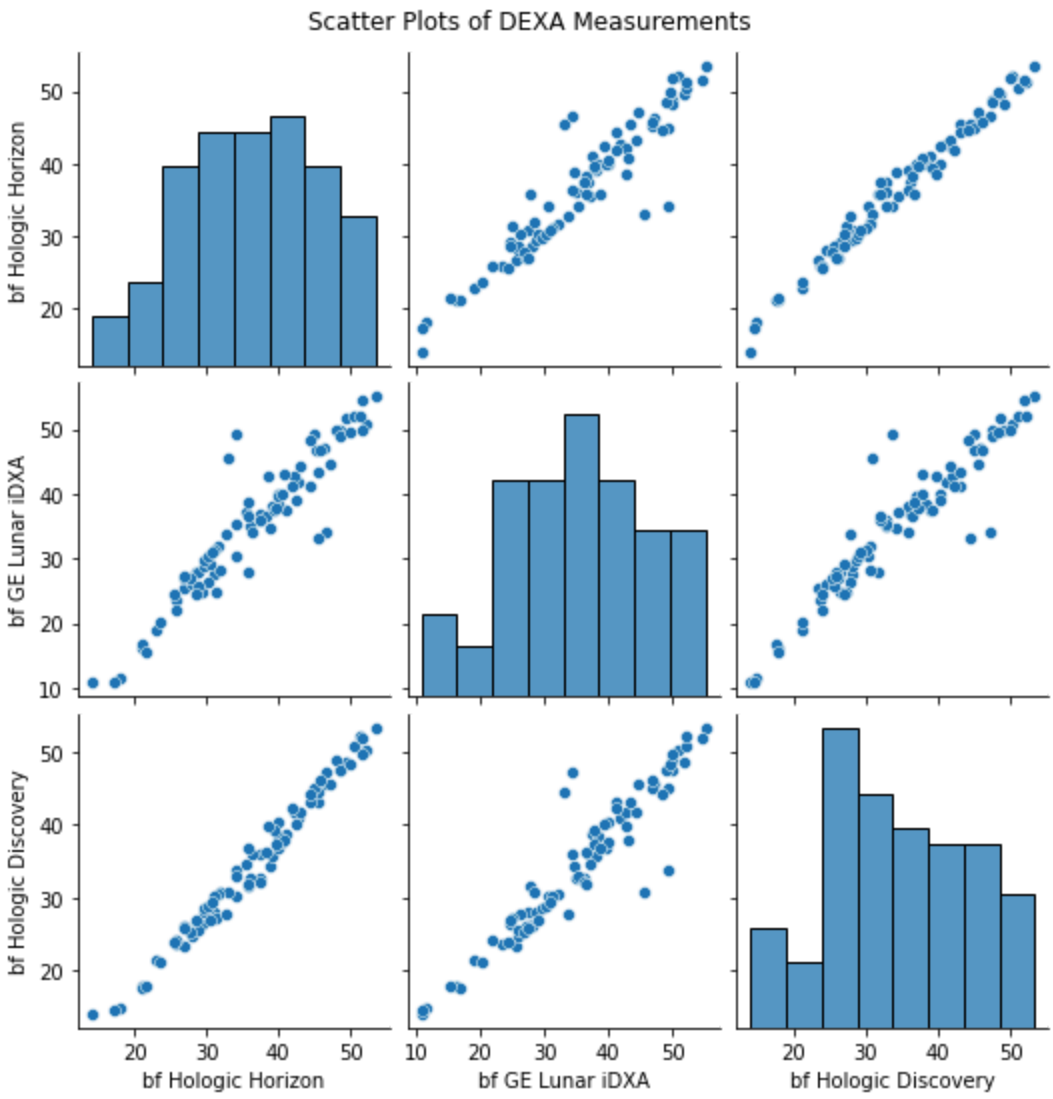
	Hologic Horizon	GE Lunar iDXA	Hologic Discovery
Hologic Horizon	1.000000	0.939067	0.990451
GE Lunar iDXA	0.939067	1.000000	0.942596
Hologic Discovery	0.990451	0.942596	1.000000

Comparison with Combined Ground-Truth
For the purposes of this study, the combined ground-truth was defined as the average body fat percentage calculated from the three DXA machines (Hologic Horizon, GE Lunar iDXA, and Hologic Discovery). This combined ground-truth provides a reference point for evaluating the accuracy of individual measurements or estimates.

The Mean Absolute Error (MAE) between the body fat percentage measurements from each DXA machine and the combined ground-truth is as follows:

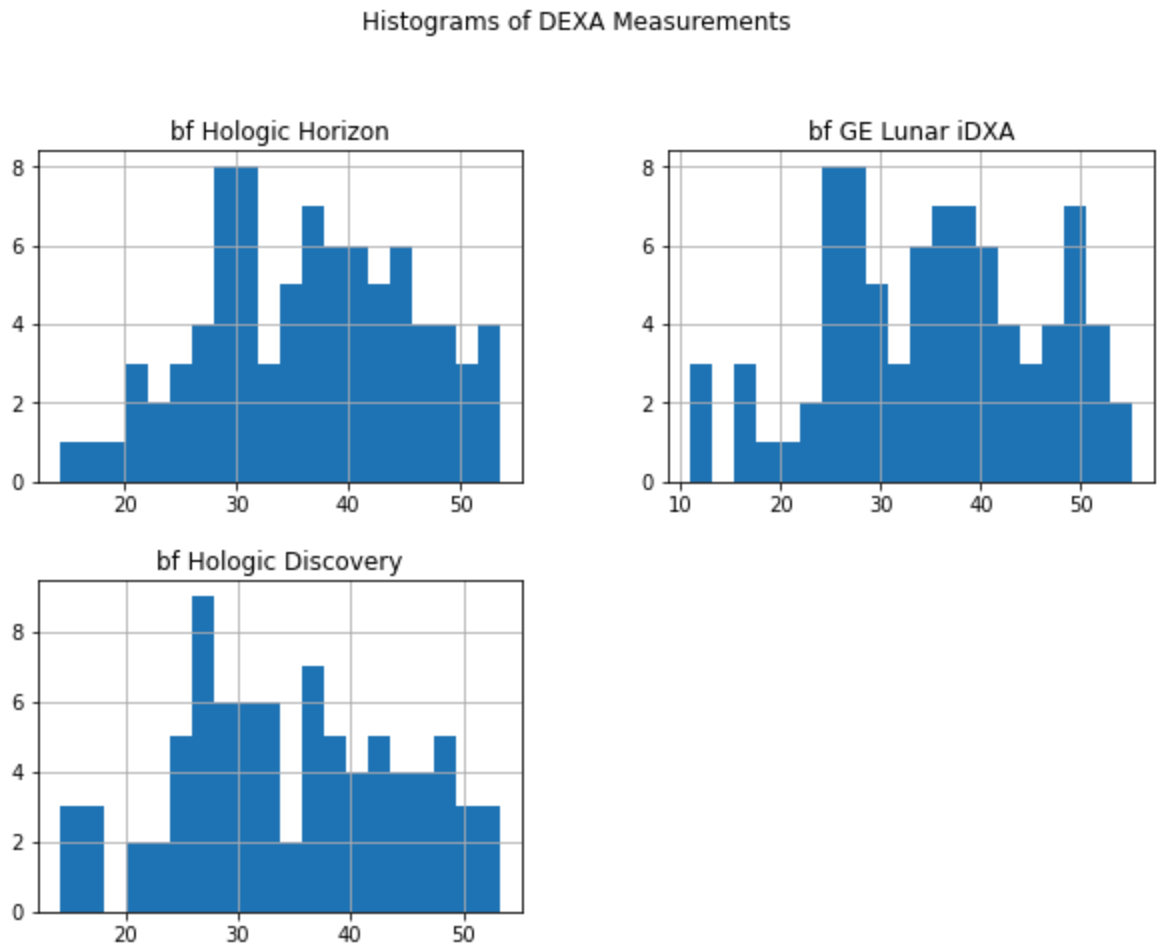
- MAE between 'Hologic Horizon' and combined ground-truth: 1.41
- MAE between 'GE Lunar iDXA' and combined ground-truth: 1.54
- MAE between 'Hologic Discovery' and combined ground-truth: 1.13

Scatter Plots of DEXA Measurements
The following plot presents scatter plots of body fat percentage measurements from the three DXA machines, illustrating the relationship and agreement between each pair of measurements.



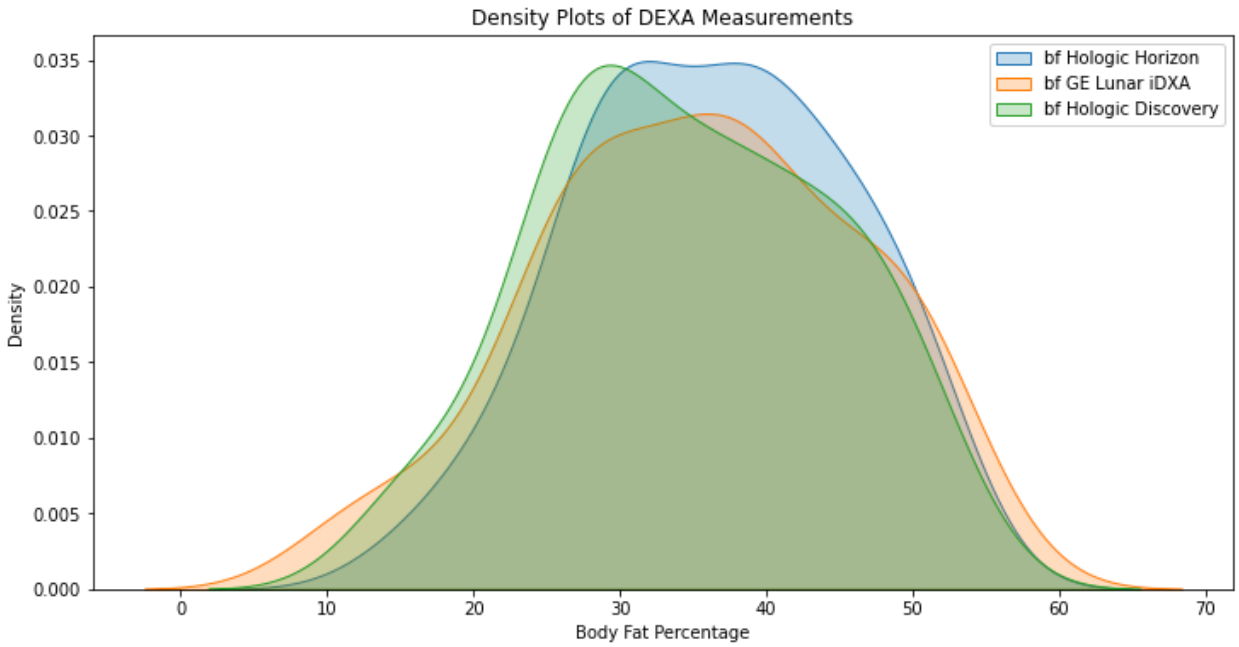
Histogram of DEXA Measurements

Histograms for body fat percentage measurements from each DXA machine provide a visual representation of the distribution of measurements.



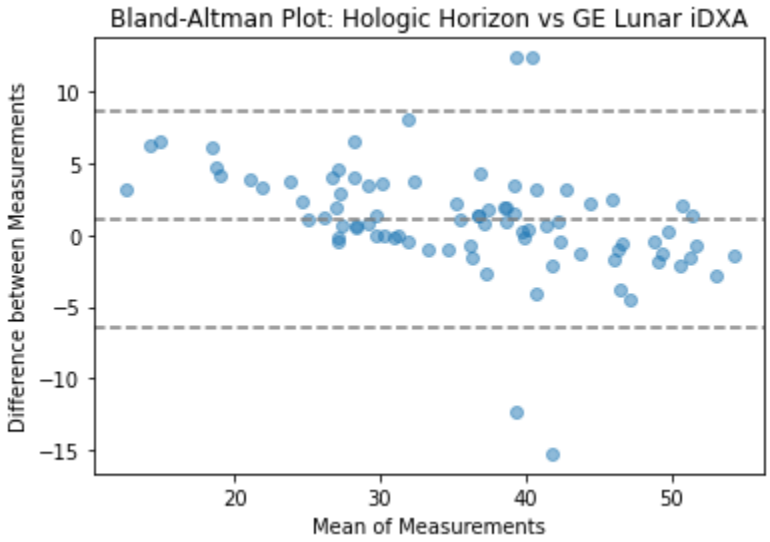
Density Plot of DEXA Measurements

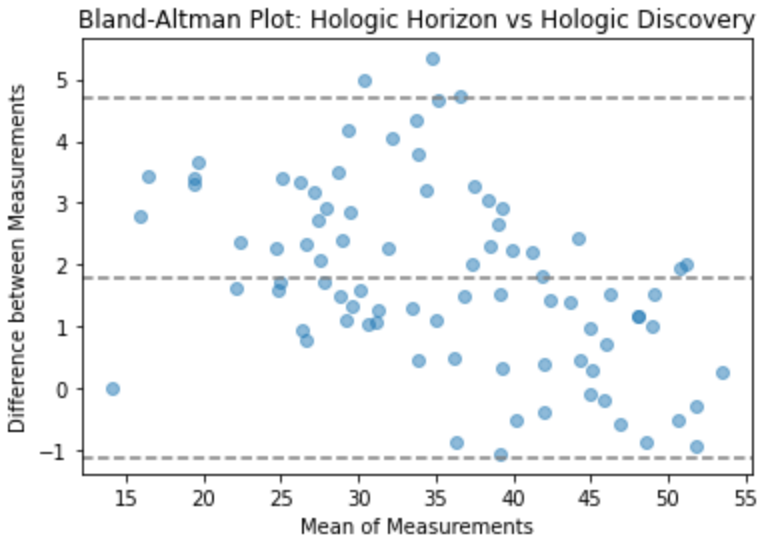
A combined density plot of the body fat percentage measurements from the three DXA machines illustrates the overlap and distribution patterns.

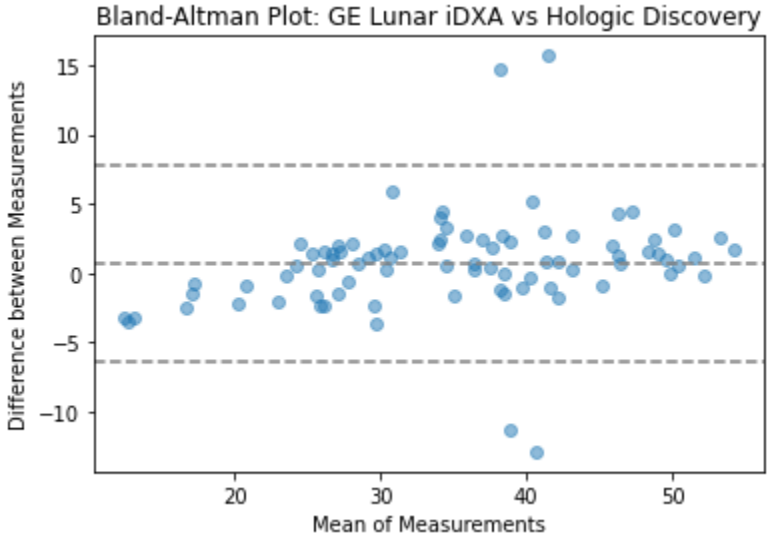


Bland-Altman Plots

Bland-Altman plots for each pair of DXA machines (Hologic Horizon vs. GE Lunar iDXA, Hologic Horizon vs. Hologic Discovery, GE Lunar iDXA vs. Hologic Discovery) provide a graphical representation of the agreement between the measurements, highlighting the mean difference and limits of agreement.

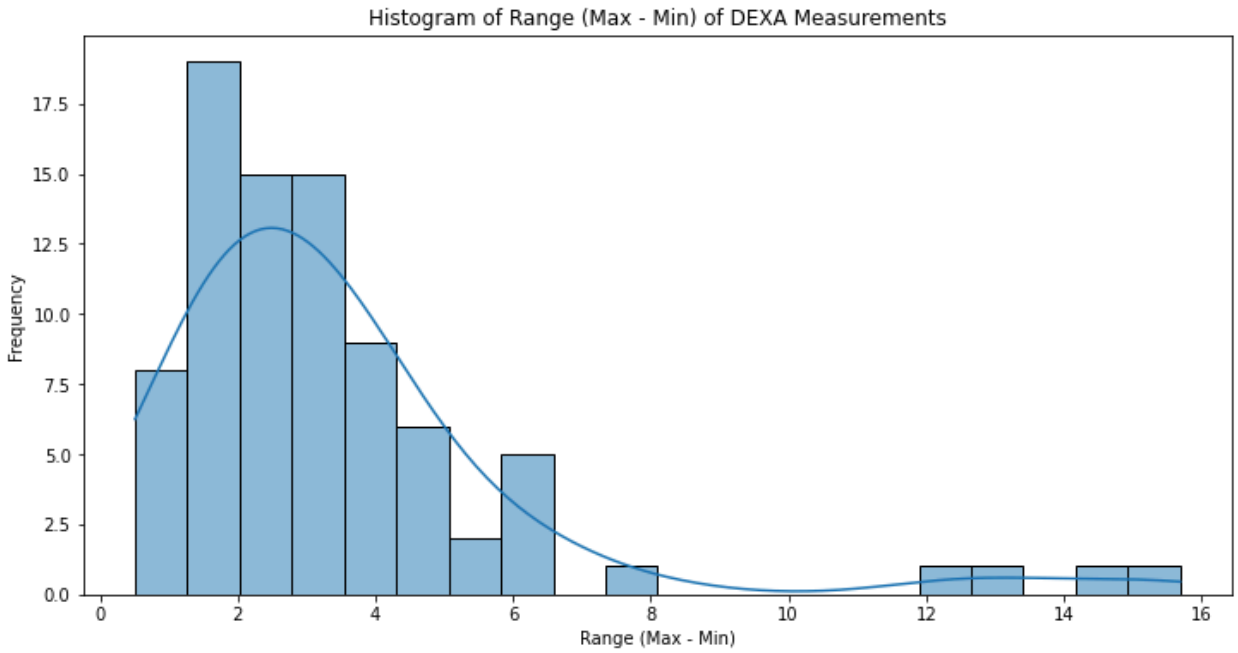






Histogram of Measurement Ranges

The histogram of the range (max - min) of body fat percentage measurements for each subject across the three DXA machines provides insights into the variability between the different machines.



Spren Prediction Comparison to DEXA Measurements

Prediction Comparison to DEXA Measurements (After Removing Subjects with High DXA Range)

For this analysis, we excluded subjects whose range of body fat percentage measurements across the three DXA machines was greater than 7%. The range here refers to the difference between the highest and lowest body fat percentage measurements recorded for a subject across the three DXA machines. This step ensures a more consistent and reliable comparison by focusing on subjects with less variability between DXA measurements. After applying this criterion, the final total number of subjects included in the analysis is 79.

Participant Demographics After Removing Subjects

- **Total number of subjects:**
  - 79
- **Gender composition:**
  - Female: 42 (53.2%)
  - Male: 37 (46.8%)
- **Ethnic and racial composition:**
  - White or Caucasian: 62 (78.5%)
  - Black or African American: 13 (16.5%)
  - South Asian: 2 (2.5%)
  - Latin American Hispanic: 2 (2.5%)

- **Mean age:** age: 50.5 ± 16.1 years (range: 23.0–78.0 years)
- **BMI:** 29.4 ± 7.0 kg/m² (range: 17.4–48.5 kg/m²)
- **DXA-measured body fat percentage (%BF):**
    - Women: 40.6 ± 8.26%
    - Men: 29.1 ± 8.41%

Mean Absolute Error (MAE) of DEXA Measurements

The Mean Absolute Error (MAE) between body fat percentage measurements from each DXA machine and the combined ground-truth, after removing subjects with high DXA range, is as follows:

- MAE between 'Hologic Horizon' and combined ground-truth: 1.23
- MAE between 'GE Lunar iDXA' and combined ground-truth: 1.14
- MAE between 'Hologic Discovery' and combined ground-truth: 0.96

The Mean Absolute Error (MAE) between body fat percentage measurements from different DXA machines is as follows:

- MAE between 'Hologic Horizon' and 'GE Lunar iDXA': 2.09
- MAE between 'Hologic Horizon' and 'Hologic Discovery': 1.96
- MAE between 'GE Lunar iDXA' and 'Hologic Discovery': 1.77

The MAE of the range (max - min) of body fat percentage across the three DXA machines is 2.90.

Summary Statistics

The following table provides summary statistics for body fat percentage predictions using Spren's app, compared to the measurements from the three DXA machines and the combined ground-truth.

Statistic	Spren Prediction	Hologic Horizon	GE Lunar iDXA	Hologic Discovery	Combined Ground-truth
Mean	33.646477	36.162813	35.092835	34.349822	35.201823
Std Dev	10.161535	9.374163	10.939352	10.040626	10.070922
Min	11.617738	14.122386	10.900700	14.122386	13.048490

25%	26.347002	29.036640	27.237100	26.703806	27.663731
50%	32.508526	36.395000	36.110900	34.204128	35.489354
75%	41.737670	42.949516	42.819800	41.968427	42.534248
Max	51.817717	53.565112	55.028700	53.298037	53.963950

Mean Absolute Error (MAE)
MAE between prediction and 'Hologic Horizon': 3.24
MAE between prediction and 'GE Lunar iDXA': 2.88
MAE between prediction and 'Hologic Discovery': 2.60
MAE between prediction and 'combined ground-truth': 2.67

Standard Deviation (STD) of Errors
STD between prediction and 'Hologic Horizon': 3.38
STD between prediction and 'GE Lunar iDXA': 3.71
STD between prediction and 'Hologic Discovery': 3.32
STD between prediction and 'combined ground-truth': 3.27

Median Absolute Error
Median AE between prediction and 'Hologic Horizon': 2.63
Median AE between prediction and 'GE Lunar iDXA': 1.93
Median AE between prediction and 'Hologic Discovery': 2.03
Median AE between prediction and 'combined ground-truth': 1.95

Pearson Correlation Coefficient
Correlation between prediction and 'Hologic Horizon': 0.94
Correlation between prediction and 'GE Lunar iDXA': 0.94
Correlation between prediction and 'Hologic Discovery': 0.95
Correlation between prediction and 'combined ground-truth': 0.95

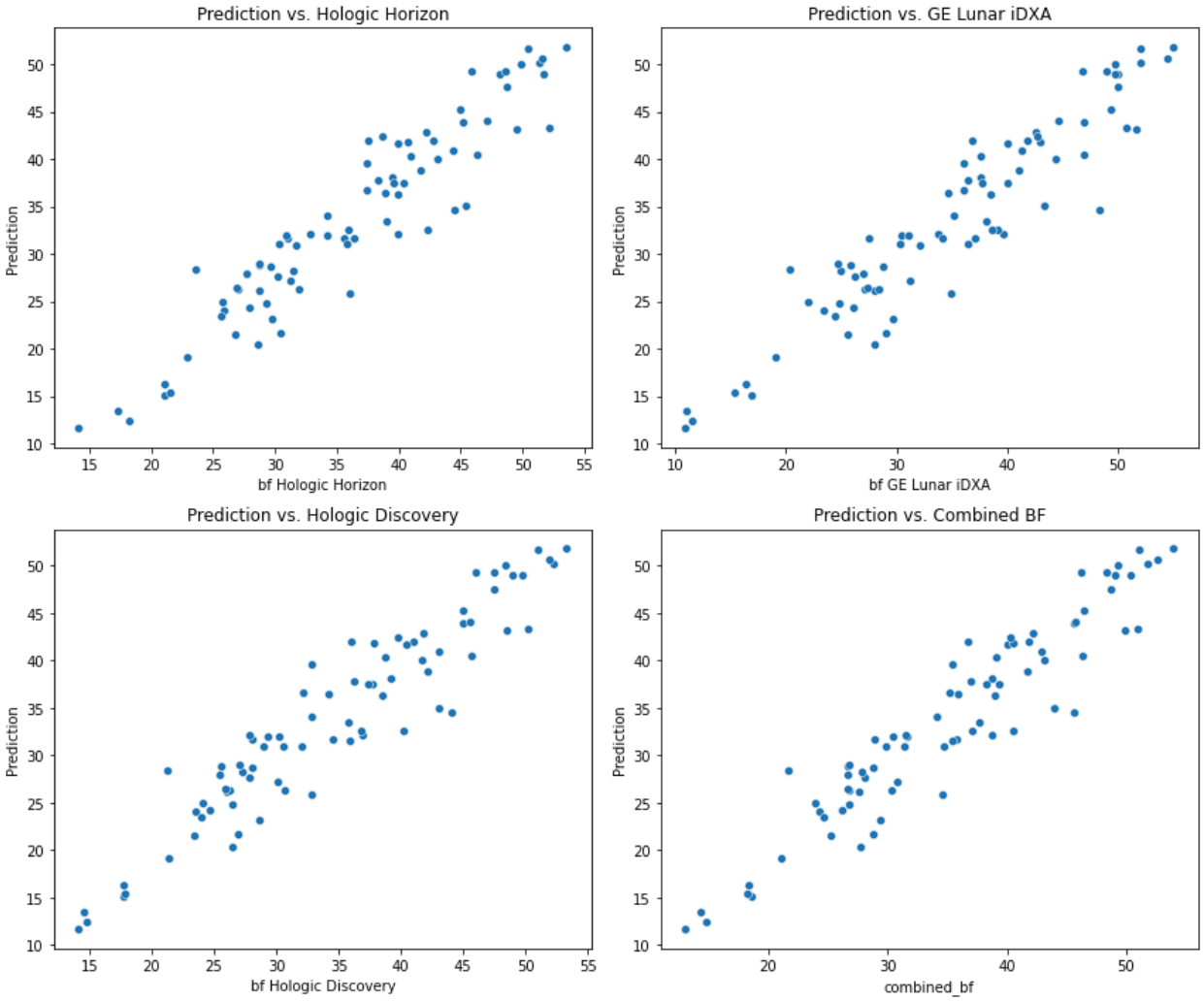
R-squared (R^2)
R^2 between prediction and 'Hologic Horizon': 0.80
R^2 between prediction and 'GE Lunar iDXA': 0.87
R^2 between prediction and 'Hologic Discovery': 0.88
R^2 between prediction and 'combined ground-truth': 0.87

Concordance Correlation Coefficient (CCC)
CCC between prediction and 'Hologic Horizon': 0.91

CCC between prediction and 'GE Lunar iDXA': 0.93
CCC between prediction and 'Hologic Discovery': 0.94
CCC between prediction and 'combined ground-truth': 0.94

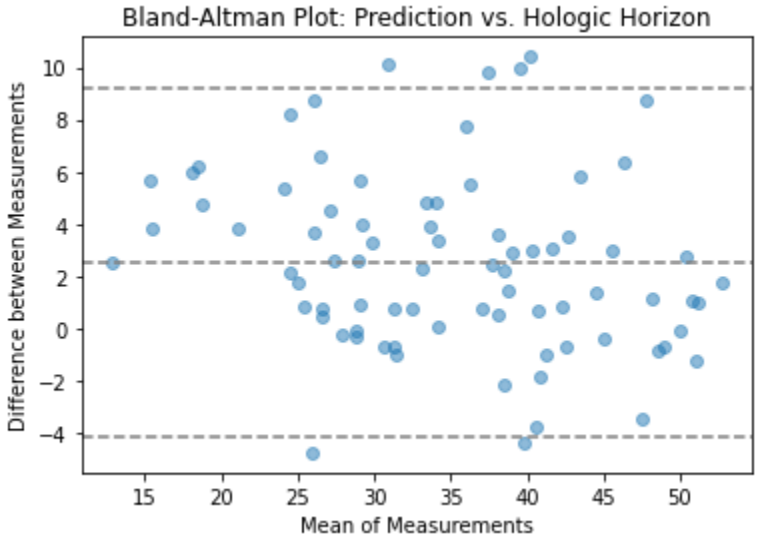
Scatter Plots of Prediction vs. DEXA Measurements

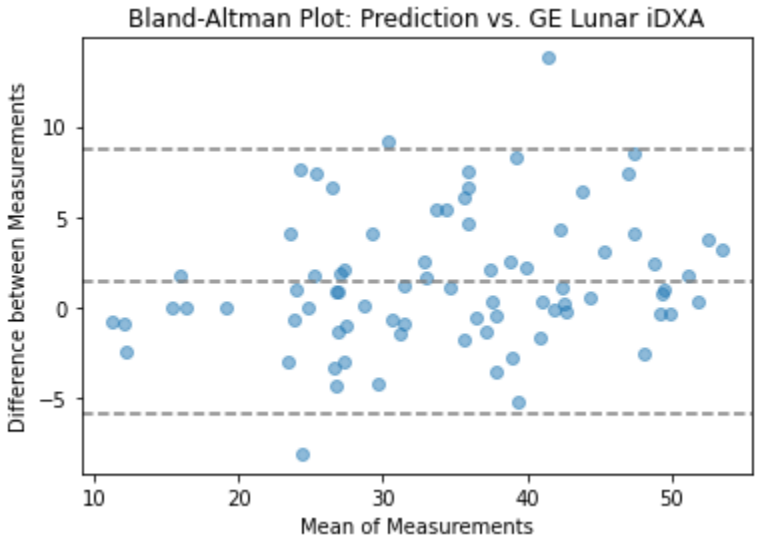
The scatter plots below compare Spren's body fat percentage predictions with the measurements from each DXA machine, illustrating the relationship and agreement between the predictions and actual measurements.

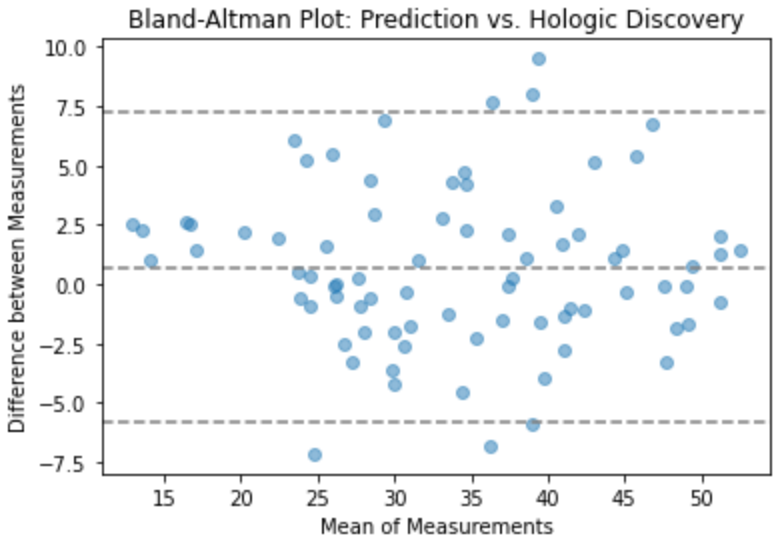


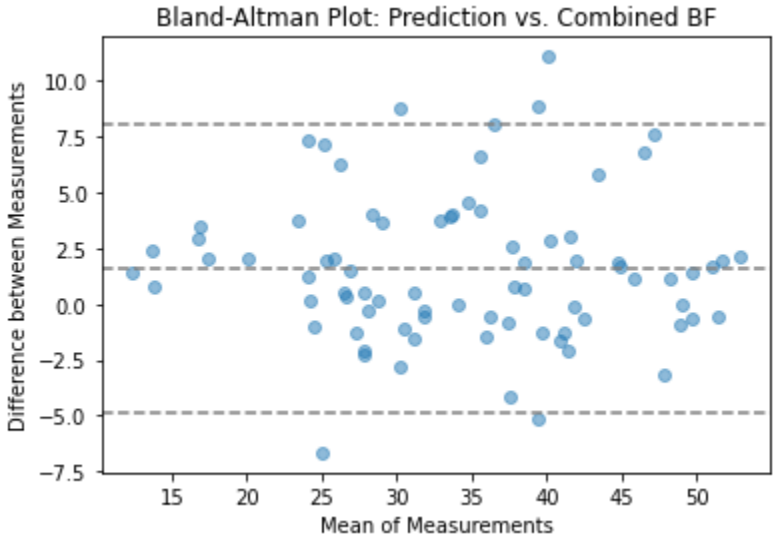
Bland-Altman Plots

The Bland-Altman plots for each pair (Spren prediction vs. each DXA machine) provide a graphical representation of the agreement between the measurements. These plots show the mean difference and limits of agreement, helping to identify any systematic bias and the range within which most differences lie.



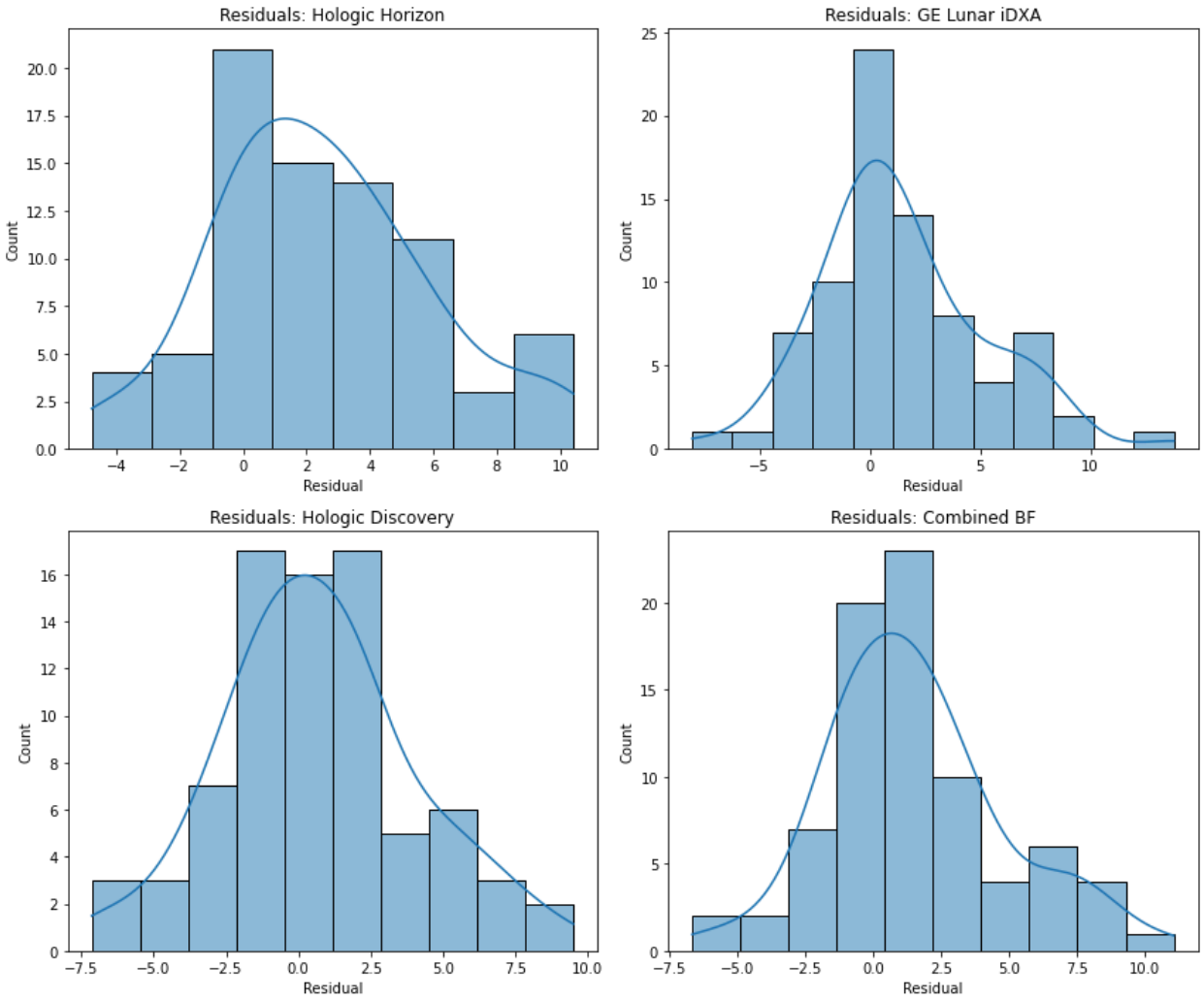






Distribution of Residuals

The distribution of residuals (prediction - actual measurements) provides insights into the errors between the Spren predictions and each DXA machine measurement, as well as the combined ground-truth. The histograms and density plots below show these distributions, highlighting the spread and central tendency of the residuals.



Correlation Matrix

The correlation matrix indicates the Pearson correlation between the body fat percentage measurements from the three DXA machines and Spren predictions.

	Spren Prediction	**Hologic Horizon**	**GE Lunar iDXA**	**Hologic Discovery**
Spren Prediction	1.000000	0.942530	0.940018	0.945373
Hologic Horizon	0.942530	1.000000	0.982663	0.990846
GE Lunar iDXA	0.940018	0.982663	1.000000	0.985052
Hologic Discovery	0.945373	0.990846	0.985052	1.000000

Performance Analysis within Sigma Ranges

To further validate the accuracy of the Spren app, we analyzed the Mean Absolute Error (MAE) and the percentage of data within one, two, and three standard deviations (sigma ranges).

Mean Absolute Error (MAE) and Percentage of Data within Sigma Ranges:

- MAE within 1 sigma (one standard deviation): 1.88
- MAE within 2 sigma (two standard deviations): 2.25
- MAE within 3 sigma (three standard deviations): 2.56

Percentage of Data within Sigma Ranges:

- 77.2% of the data falls within 1 sigma (one standard deviation).
- 93.6% of the data falls within 2 sigma (two standard deviations).
- 98.7% of the data falls within 3 sigma (three standard deviations).

Performance Across Demographics

The performance of Spren's body fat percentage predictions was also evaluated across different demographic groups.

MAE by Gender

Gender	MAE
Female	2.372721
Male	3.017066

MAE by Race

Race	MAE
Black or African American	1.343601
Latin American Hispanic	1.989766

South Asian	2.509581
White or Caucasian	2.980971

MAE by BMI Categories

BMI Category	MAE
< 18.5	1.368299
18.5--25	2.555631
25-30	3.122235
> 30	2.583410

Repeatability of Spren Predictions

To assess the repeatability of Spren's body fat percentage predictions, we compared the two scans taken for each subject. The following results demonstrate the discrepancies (defined as the absolute difference) between the two predictions and the consistency of these predictions.

Summary Statistics of Discrepancies

Statistic	Value
Mean	0.999104
Std Dev	0.969117
Min	0.021470
25%	0.312498
50%	0.678792

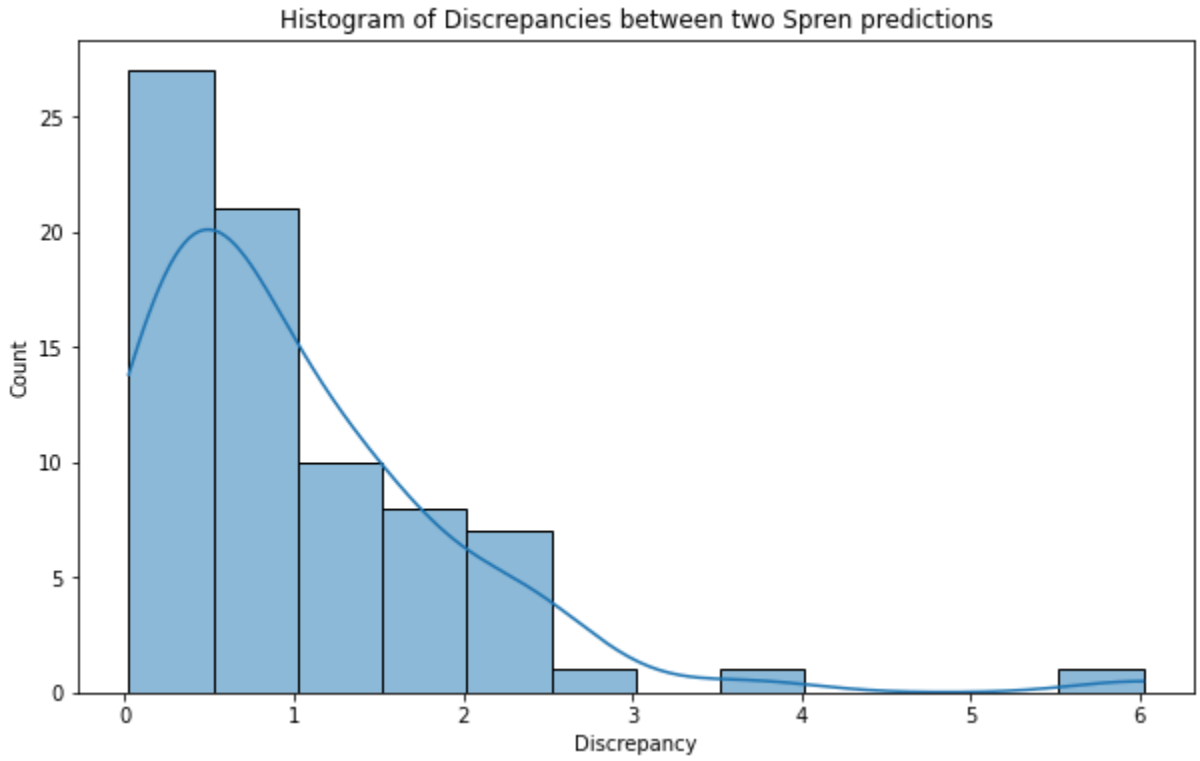
75%	1.456969
Max	6.022805

Additional Metrics

- Mean Absolute Error (MAE): 1.00
- Median Absolute Error: 0.68
- Standard Deviation: 0.96
- Pearson Correlation Coefficient: 0.99
- R-squared (R^2): 0.98
- Concordance Correlation Coefficient (CCC): 0.99

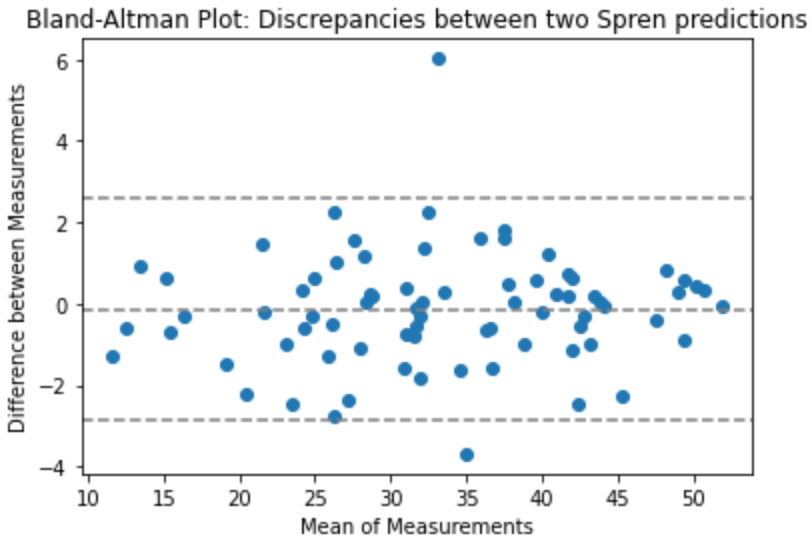
Distribution of Discrepancies

The distribution of discrepancies (defined as the absolute difference) between the two predictions.



Bland-Altman Plots

The Bland-Altman plots for discrepancies provide a graphical representation of the agreement between the measurements.



Conclusion

The findings of this validation study highlight the accuracy and reliability of Spren's body composition app in estimating body fat percentage using whole body photos taken by smartphone camera. Our comprehensive analysis, involving a diverse participant pool and comparison with gold-standard DXA measurements, demonstrates that the app achieves a mean absolute error (MAE) of 2.6% and a Pearson correlation of 0.95 with the combined DXA ground-truth. These results indicate a high degree of concordance and suggest that Spren's app offers a viable, non-invasive alternative for body composition assessment.

Furthermore, the performance analysis within sigma ranges validates the app's accuracy, with 77.2% of predictions falling within one standard deviation (1 sigma) with an MAE of 1.88, 93.6% within two standard deviations (2 sigma) with an MAE of 2.25, and 98.7% within three standard deviations (3 sigma) with an MAE of 2.56 of the actual measurements.

Moreover, the app's repeatability was confirmed through dual scanning, showing minimal variability between repeated measurements, with a Mean Absolute Error (MAE) of 1.00 and a Pearson correlation coefficient of 0.99 between the two scans.

Our demographic analysis indicates that Spren's predictions are consistent across different genders, races, and BMI categories, although minor variations were observed. Specifically, the MAE was 2.37 for females and 3.01 for males. By race, the MAE was 1.34 for Black or African American participants, 1.98 for Latin American Hispanic participants, 2.51 for South Asian participants, and 2.98 for White or Caucasian participants. For BMI categories, the MAE was 1.36 for participants with a BMI less than 18.5, 2.55 for those with a BMI between 18.5 and 25, 3.12 for those with a BMI between 25 and 30, and 2.58 for those with a BMI over 30. This robustness across diverse populations enhances the app's potential for widespread use in public health monitoring and personalized health management.

When comparing the distribution of body fat percentage estimates from the Spren app to the combined ground-truth distribution from the DXA measurements, the statistical analysis revealed that the mean body fat percentage predicted by the Spren app closely matches the combined ground-truth mean. The standard deviation of the Spren app predictions is comparable to that of the combined ground-truth, and the minimum and maximum values of the Spren predictions fall within the range of the combined ground-truth values. These findings suggest that the Spren app not only provides accurate individual estimates but also reflects the overall distribution and variability of body fat percentages in the population, similar to the ground-truth measurements obtained from DXA scans.

In conclusion, Spren's body composition app presents a reliable, accessible, and cost-effective method for estimating body fat percentage, with significant implications for public health. Its non-invasive nature and ease of use make it an attractive alternative to traditional DXA scans, potentially democratizing body composition assessment and facilitating better health and fitness monitoring on a large scale. Future research could focus on further refining the app's algorithm and expanding its application to other health metrics, thereby broadening its impact in the healthcare industry.