EAST WEST UNIVERSITY

EDA And Related Work Report(Task-1)
Group-9

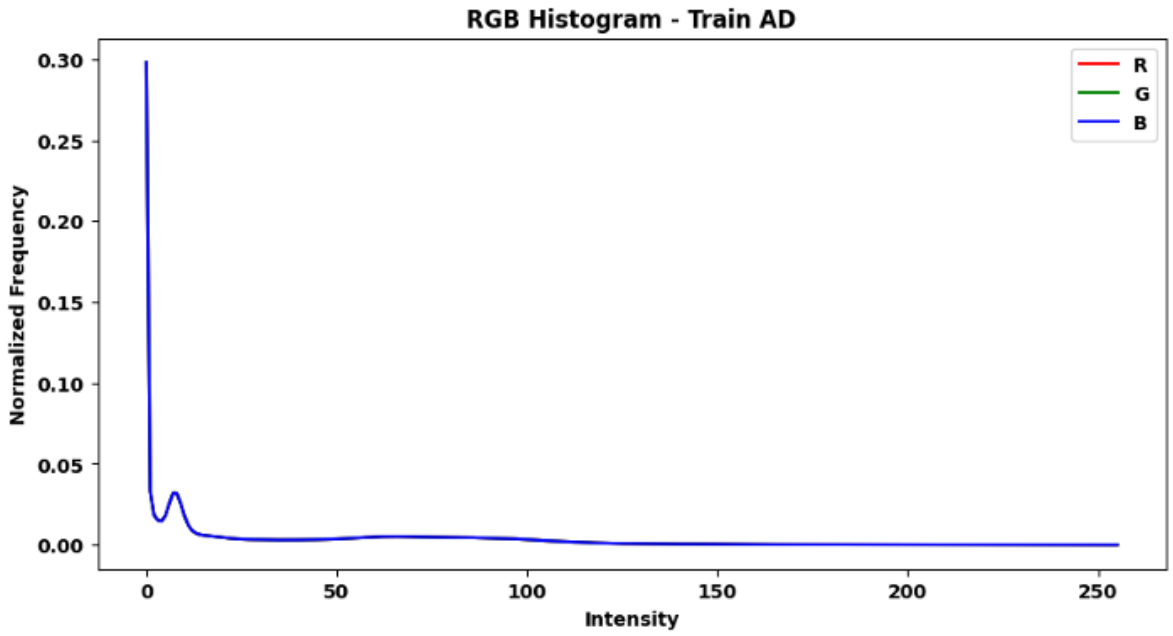
Group Members

Name	ID
Farhan Ibtesham Joy	2022-3-60-150
Miftha Ahona Majid	2022-1-50-015
Mithila Akter Aka	2022-1-50-010

Submitted To:

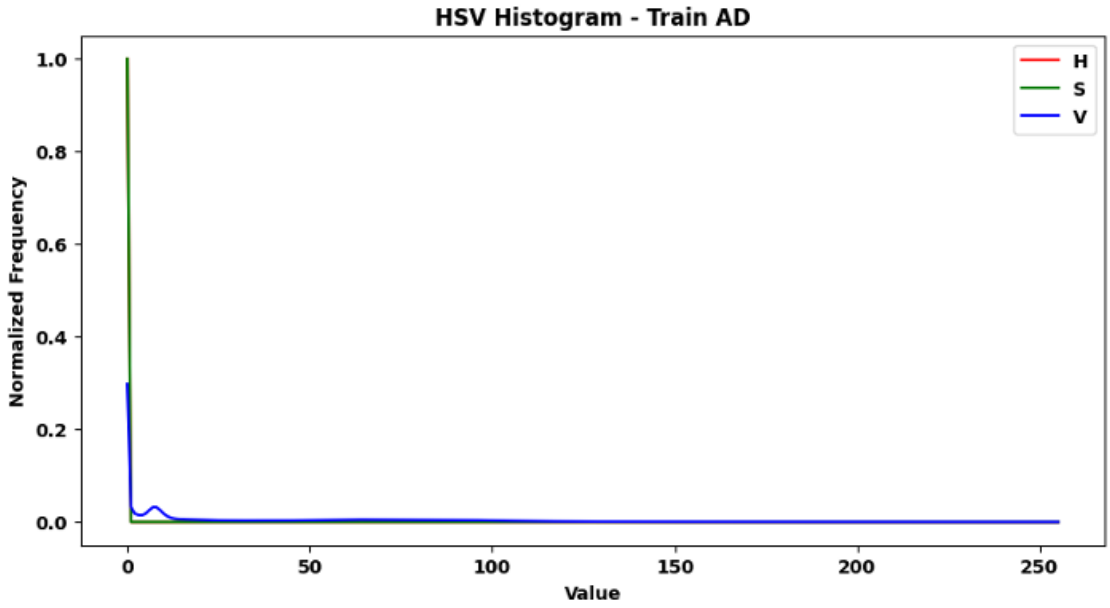
Dr. Raihan Ul Islam
Associate Professor
Dept. Of Computer Science & Engineering
East West University

Figure-1



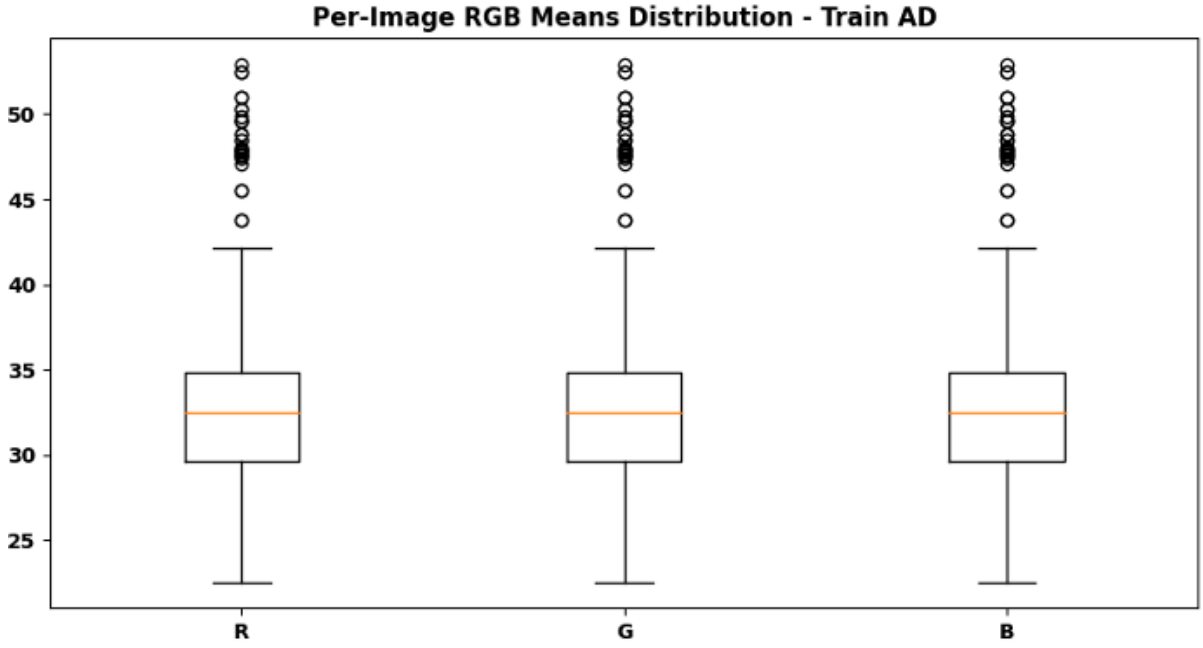
This RGB Histogram Train AD (Test) Displays sample Alzheimer's Disease test images; shows grayscale MRI slices illustrating structural brain differences.Shows sample Alzheimer's Disease test images.It displays grayscale MRI brain scans with visible structural differences.
These images are used to detect Alzheimer's patterns.

Figure-2



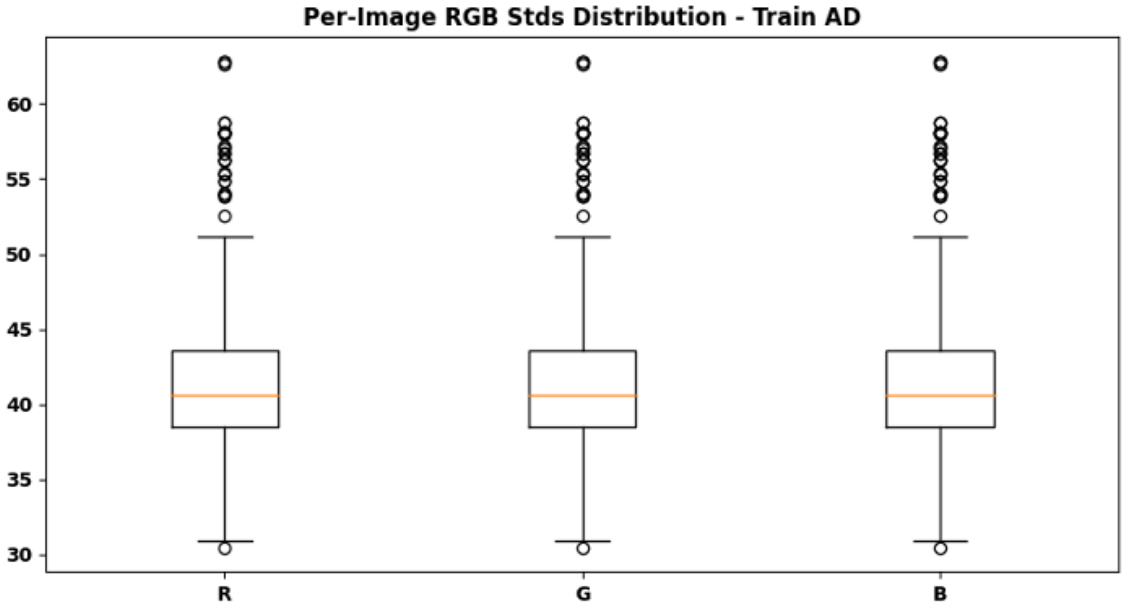
Shows Control Normal (CN) test images.These brains appear normal without Alzheimer's effects.Used for comparing healthy and affected brains.

Figure-3



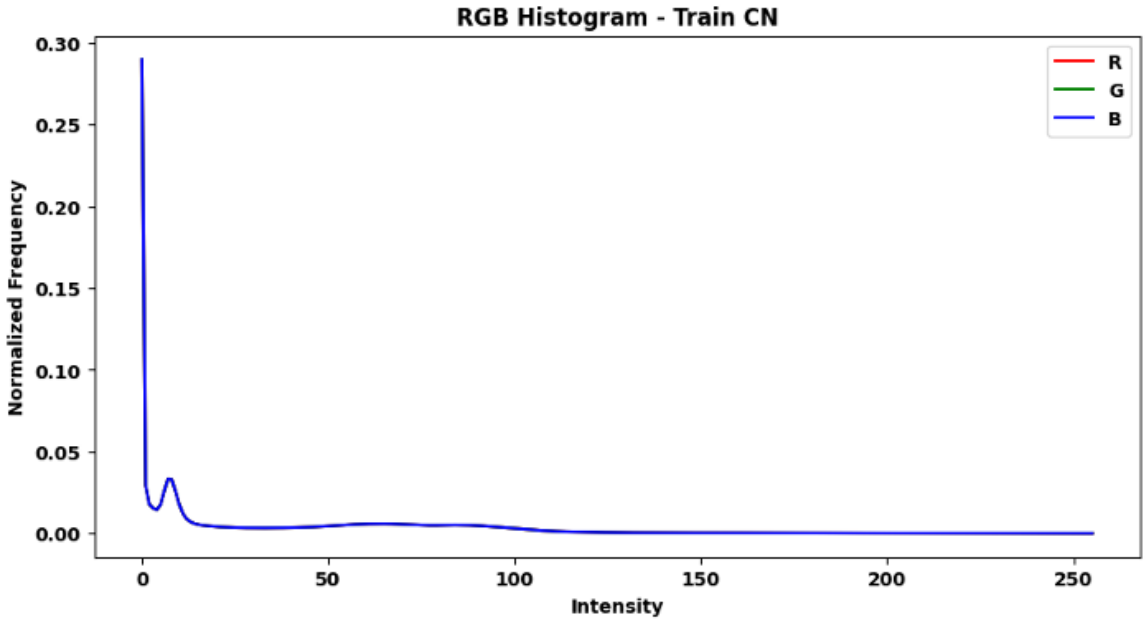
Displays Mild Cognitive Impairment (MCI) test images.These images show early brain changes.They help in detecting the transition from normal to Alzheimer's.

Figure-4



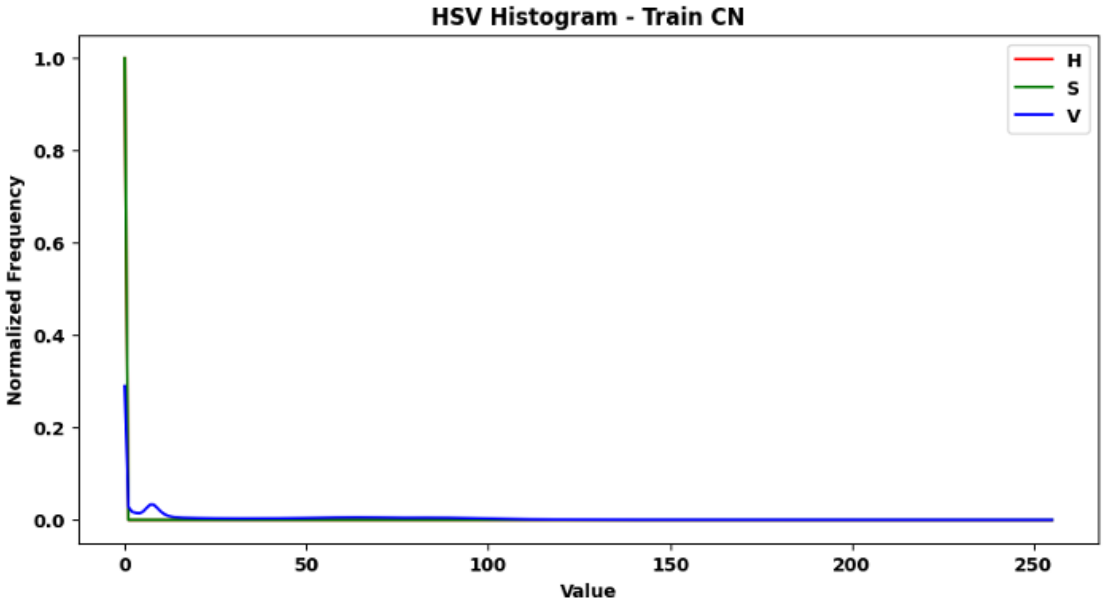
Shows Alzheimer's Disease (AD) training images.Used to teach the model Alzheimer's features.The patterns help the model learn memory-related brain loss.

figure -5



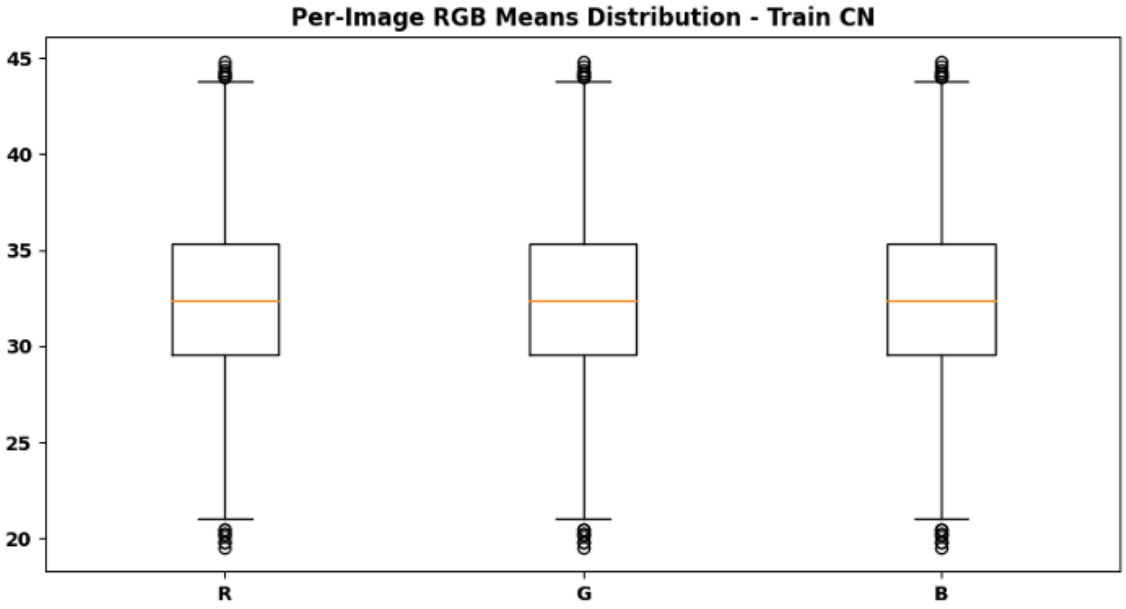
Displays CN (Control Normal) training images.These help the system learn what healthy brains look like.They are important for comparison with AD cases.

Figure-6



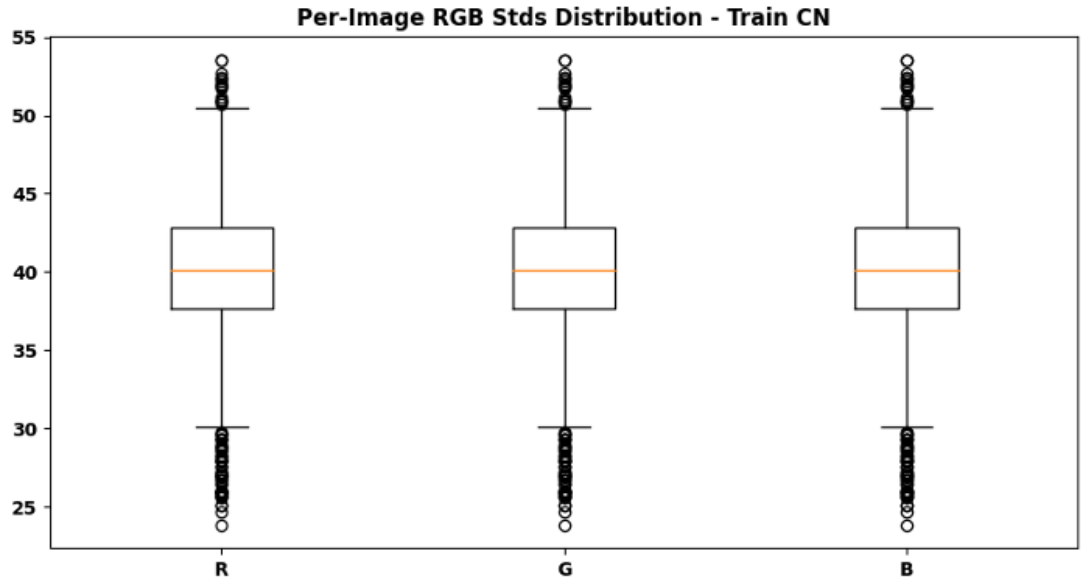
Shows MCI (Mild Cognitive Impairment) training images.It shows small brain differences between normal and AD.These images improve early detection accuracy.

Figure-7



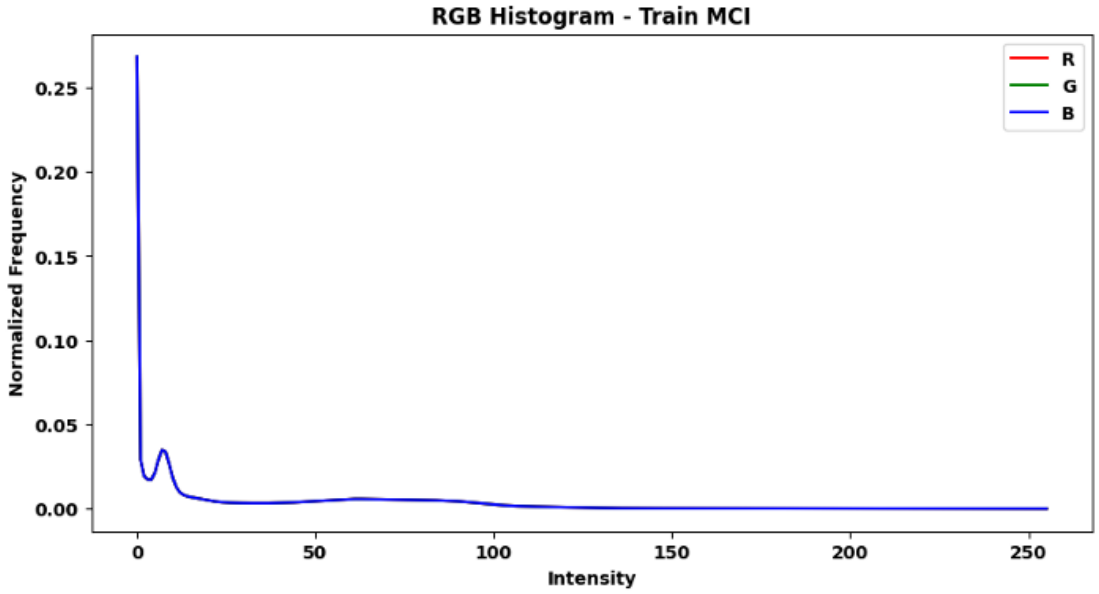
Shows GPU detection and setup results.It confirms that two GPUs are working properly.This ensures faster training and better performance.

Figure-8



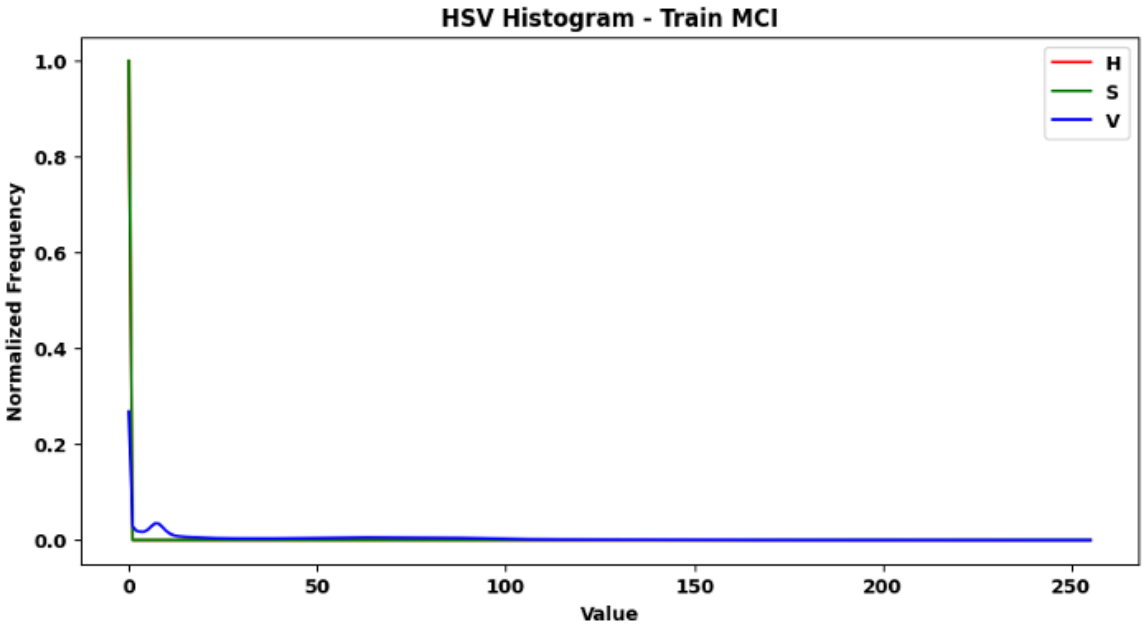
Displays RGB histogram for Alzheimer's training images.It shows how color values (R, G, B) are distributed.Indicates balanced intensity across channels.

Figure-9



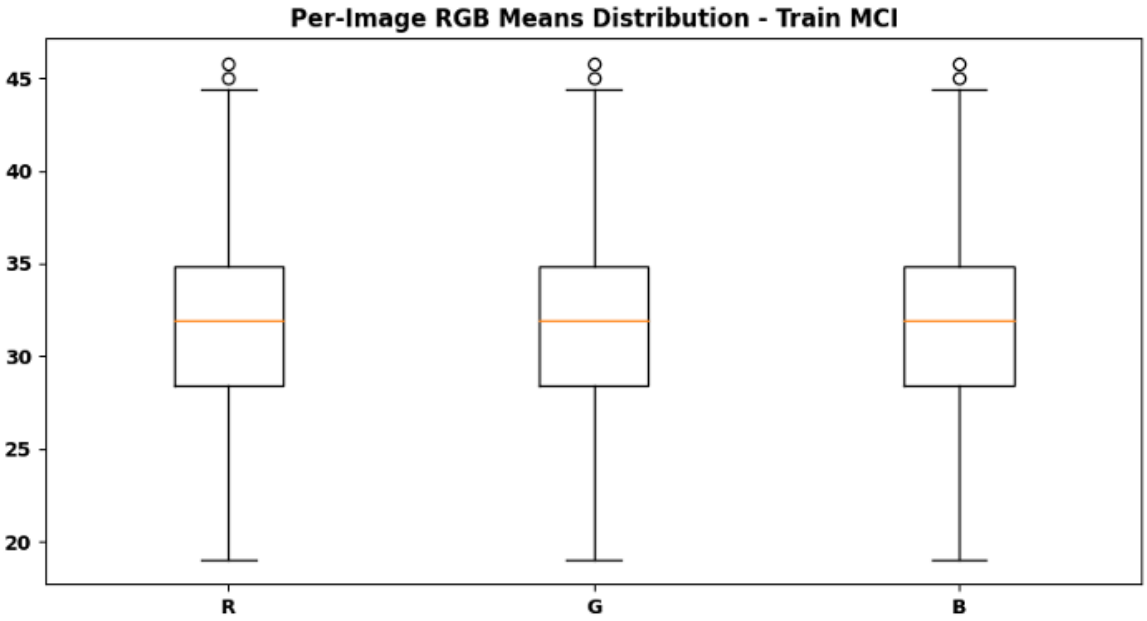
Shows HSV (Hue, Saturation, Value) histogram for AD training data.It represents color tone, brightness, and variation.Helps check the visual quality of images.

Figure-10



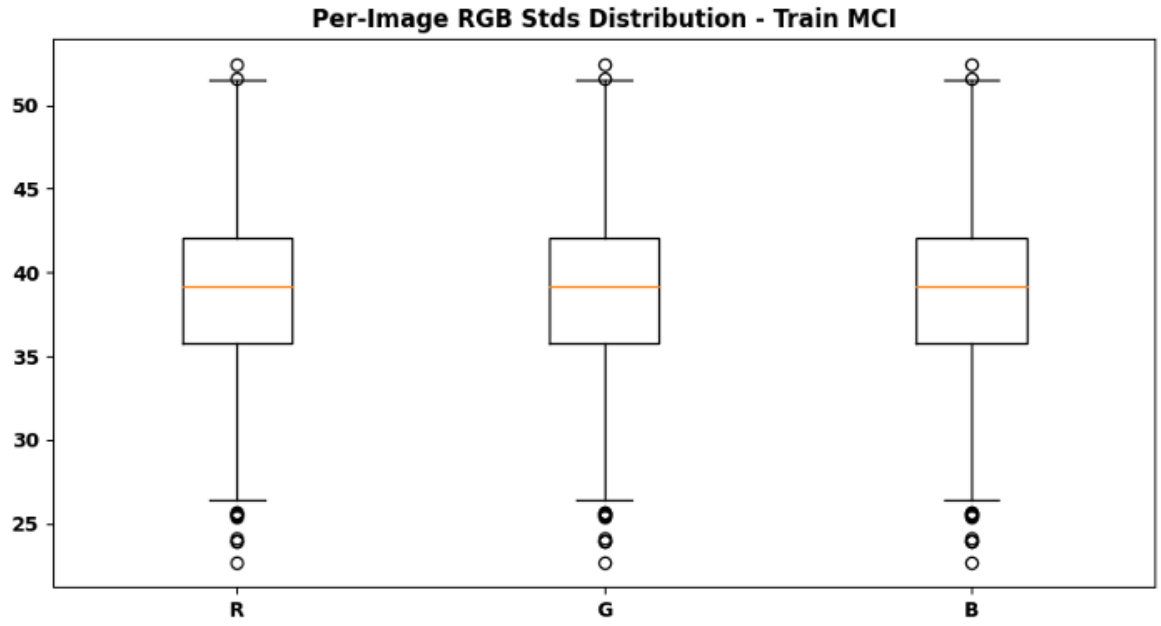
Displays RGB histogram for Control Normal training data.Shows that color balance is similar to AD data.Confirms consistent image preprocessing.

Figure-11



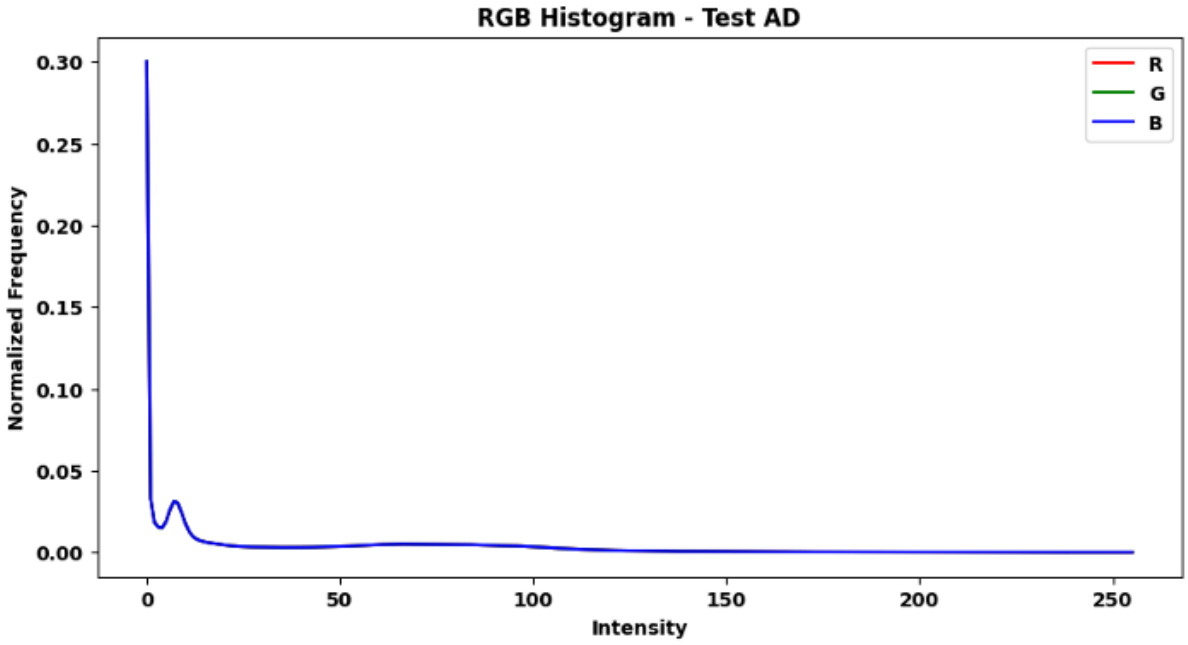
Shows HSV histogram for CN training images.Brightness and color tones are stable.Confirms good dataset lighting conditions.

Figure-12



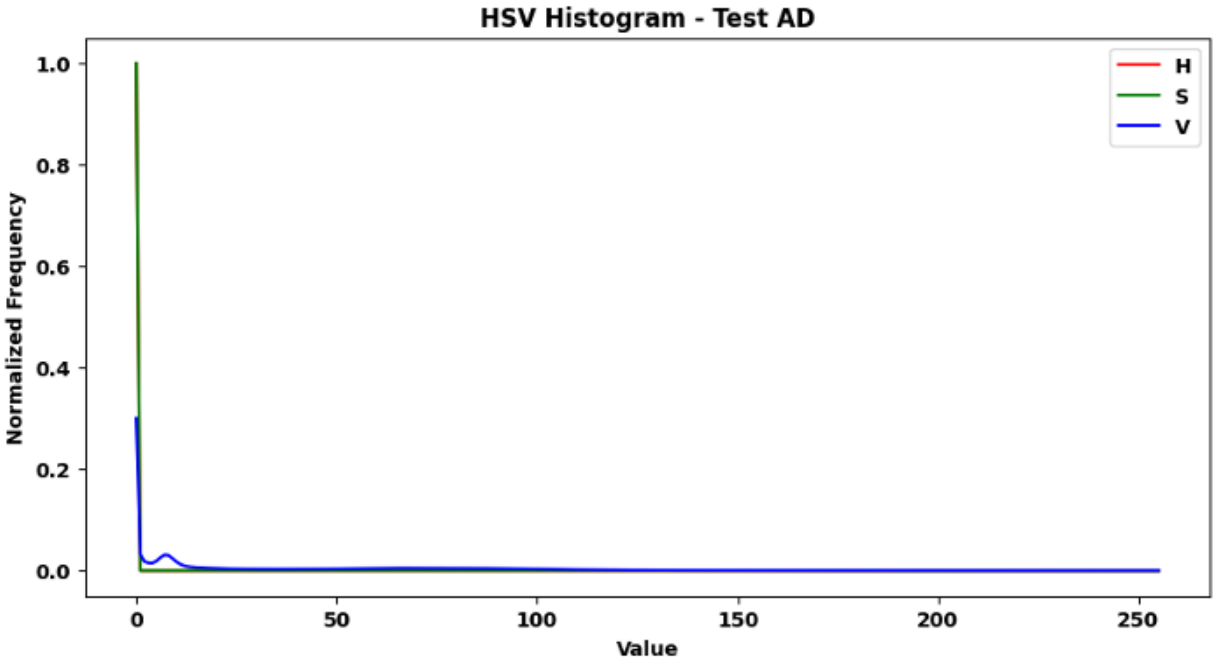
Displays RGB histogram for MCI training images.Shows slightly darker intensity levels.Indicates mild variation in color range.

Figure-13



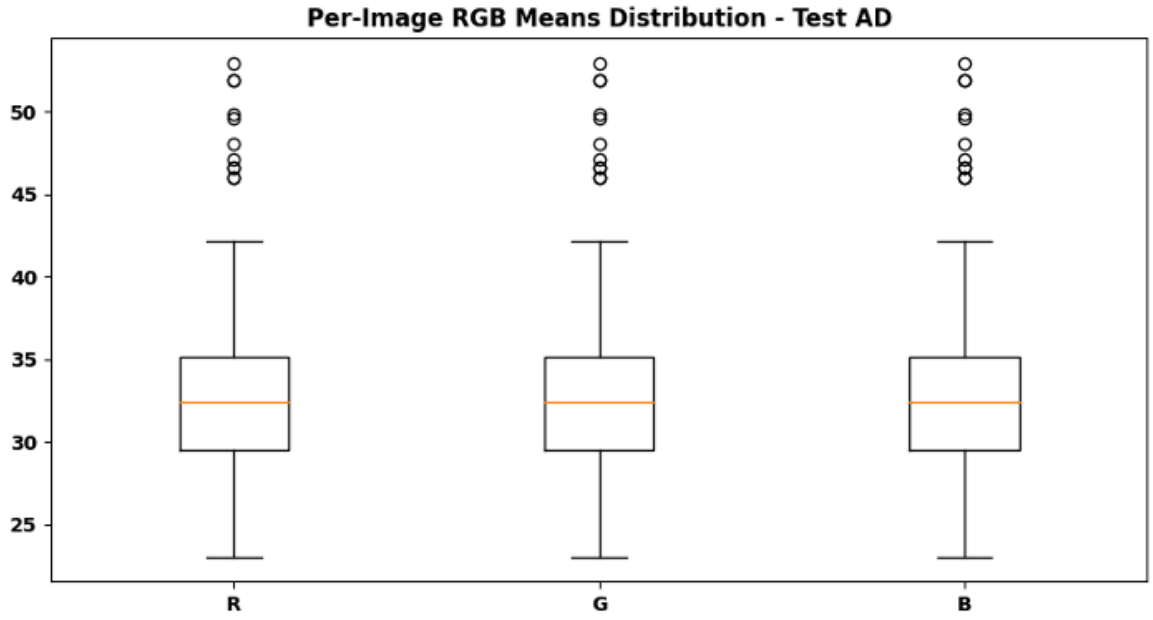
Shows HSV histogram for MCI training data.Brightness is uniform with low contrast.Useful for checking image consistency.

Figure-14



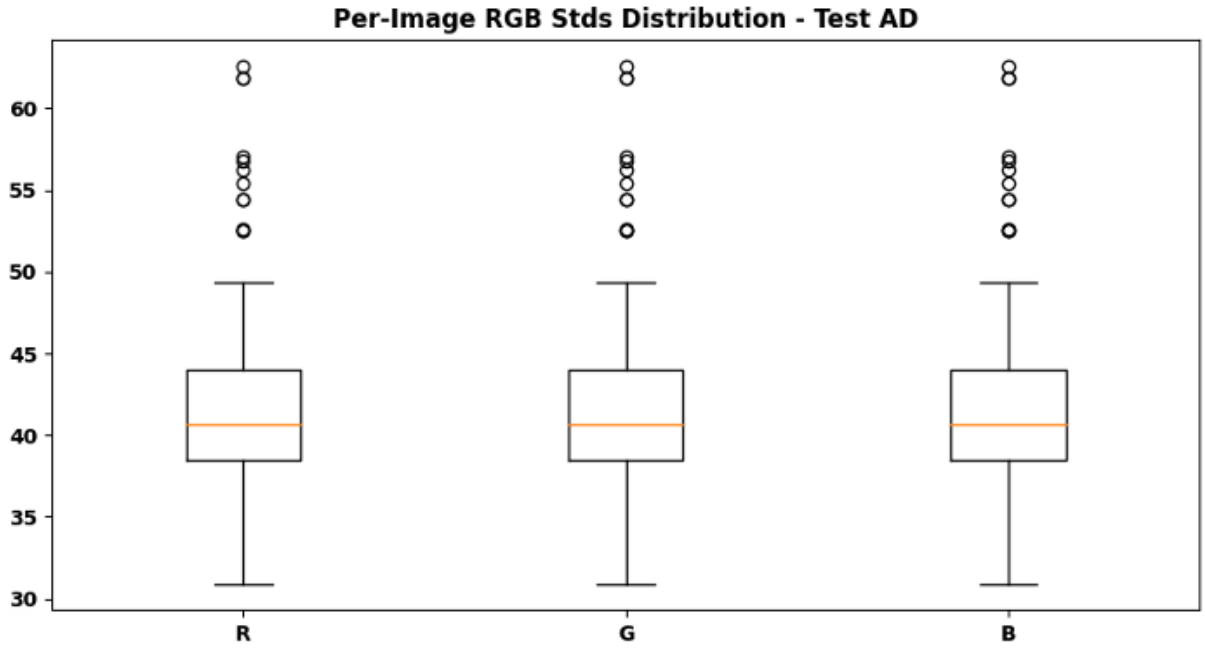
Shows RGB histogram for Alzheimer's test images.Color pattern similar to training data.Ensures no major difference between sets.

Figure-15



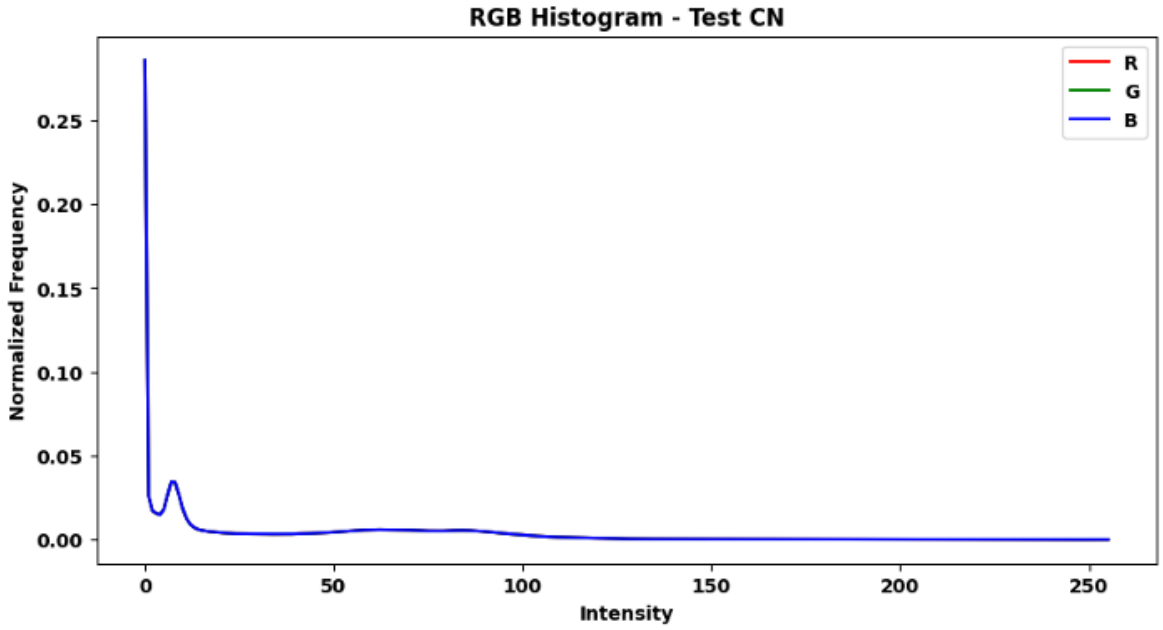
Displays HSV histogram for AD test images.Shows consistent brightness levels.Indicates test data quality is good.

Figure-16



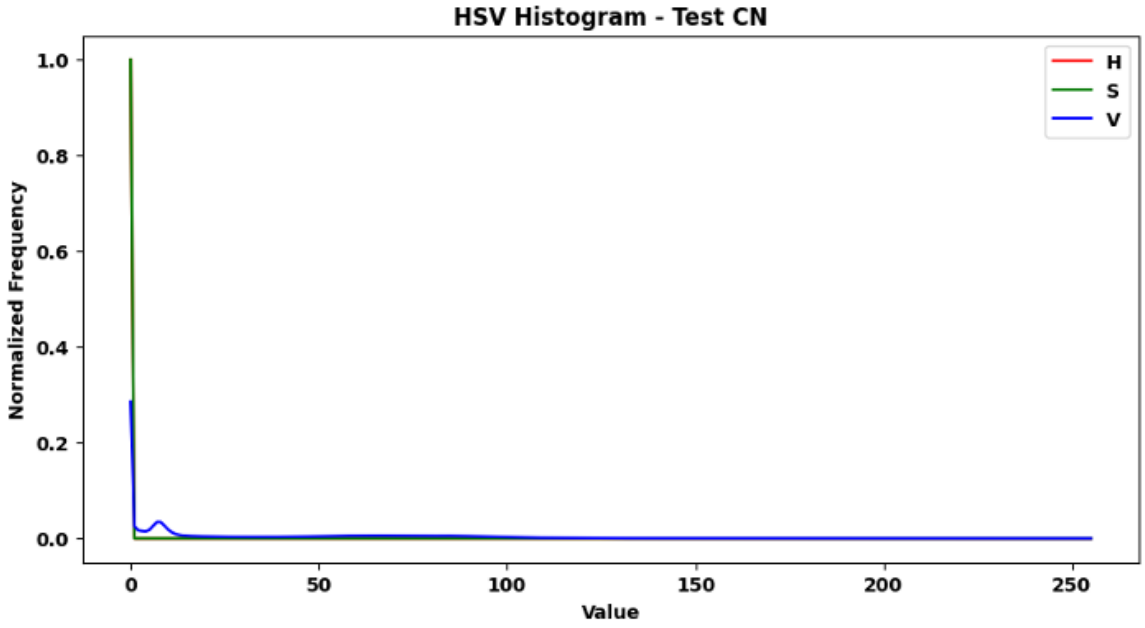
Shows RGB histogram for CN test data.All three color channels are well balanced.Confirms even image brightness.

Figure-17



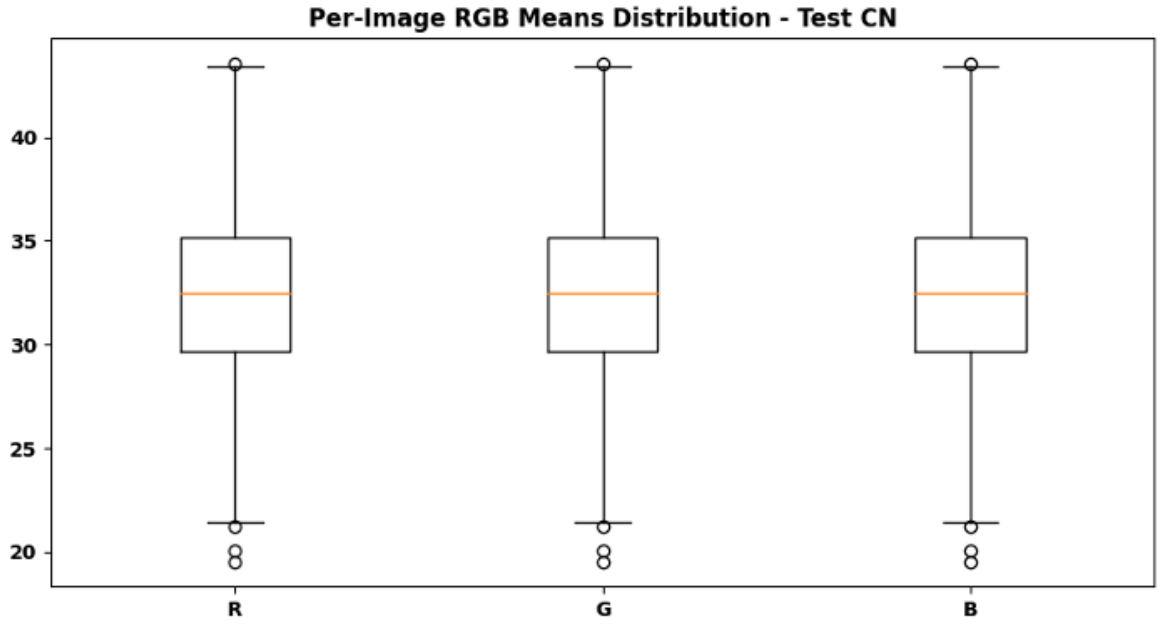
Displays HSV histogram for CN test images.Shows stable brightness and color tones.Indicates uniform lighting.

Figure-18



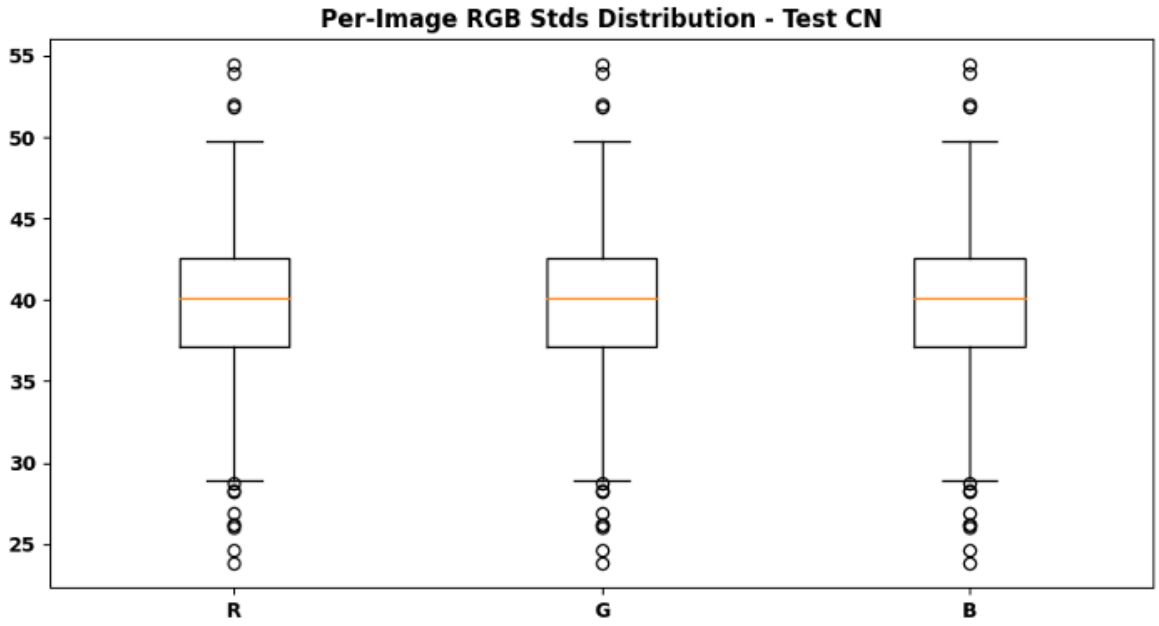
Shows RGB histogram for MCI test data.Color intensity is slightly low.Images remain consistent with training set.

Figure-19



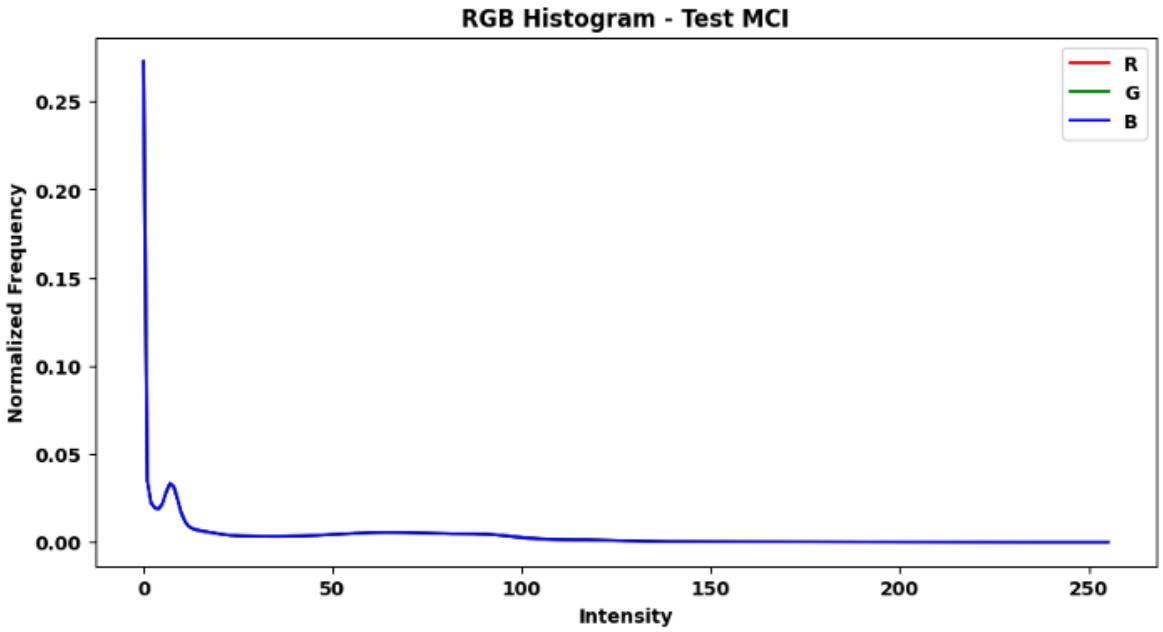
Displays HSV histogram for MCI test images.Shows smooth brightness distribution.Proves equal exposure across images.

Figure-20



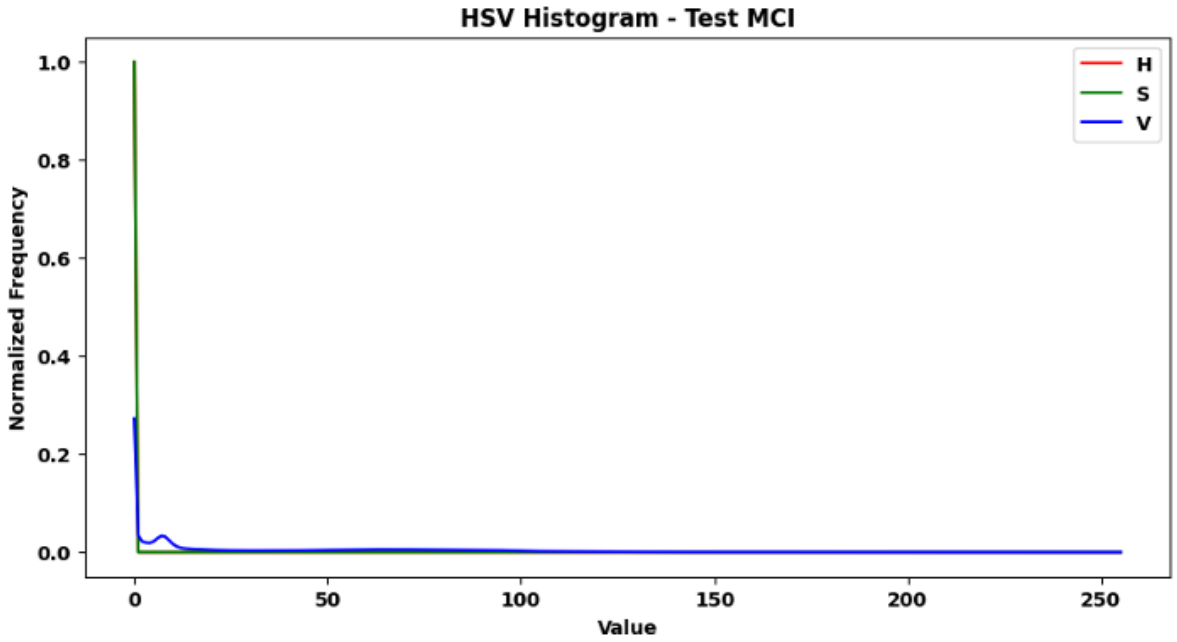
Shows per-image RGB mean values for AD training data.Each box shows how bright images are.Most images have similar color intensity.

Figure-21



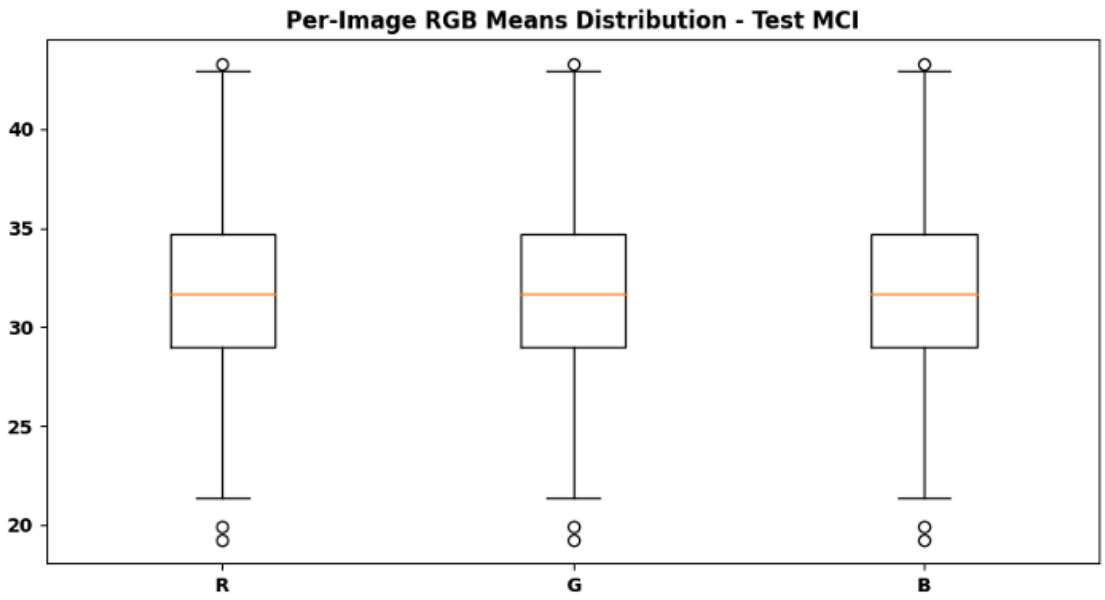
Displays per-image RGB standard deviation for AD training data.Shows how much brightness varies between images.Low variation means dataset is uniform.

Figure-22



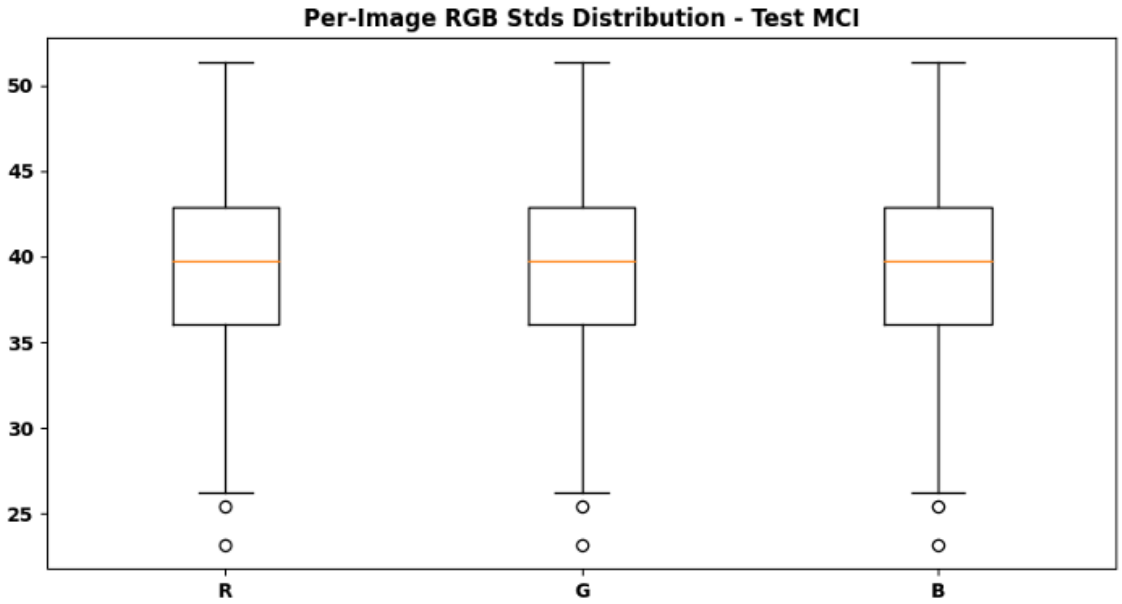
Shows per-image RGB mean for CN training data.All three color channels have close values.Indicates consistent imaging process.

Figure-23



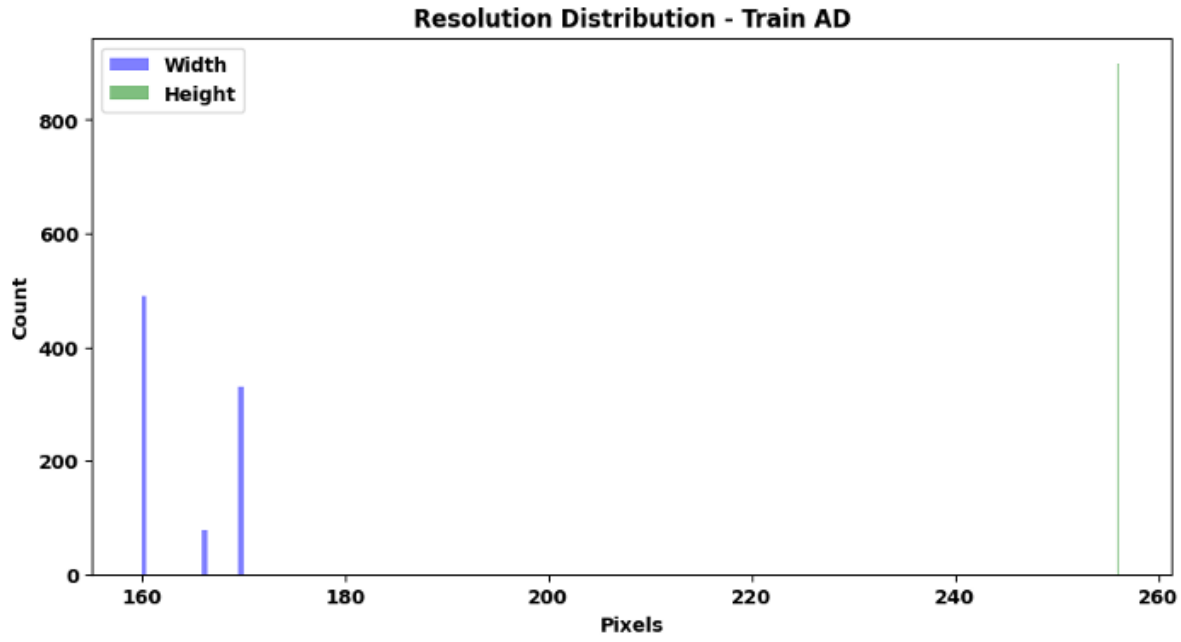
Displays RGB standard deviation for CN training images.Shows stable color and contrast.Helps confirm clean, balanced dataset.

Figure-24



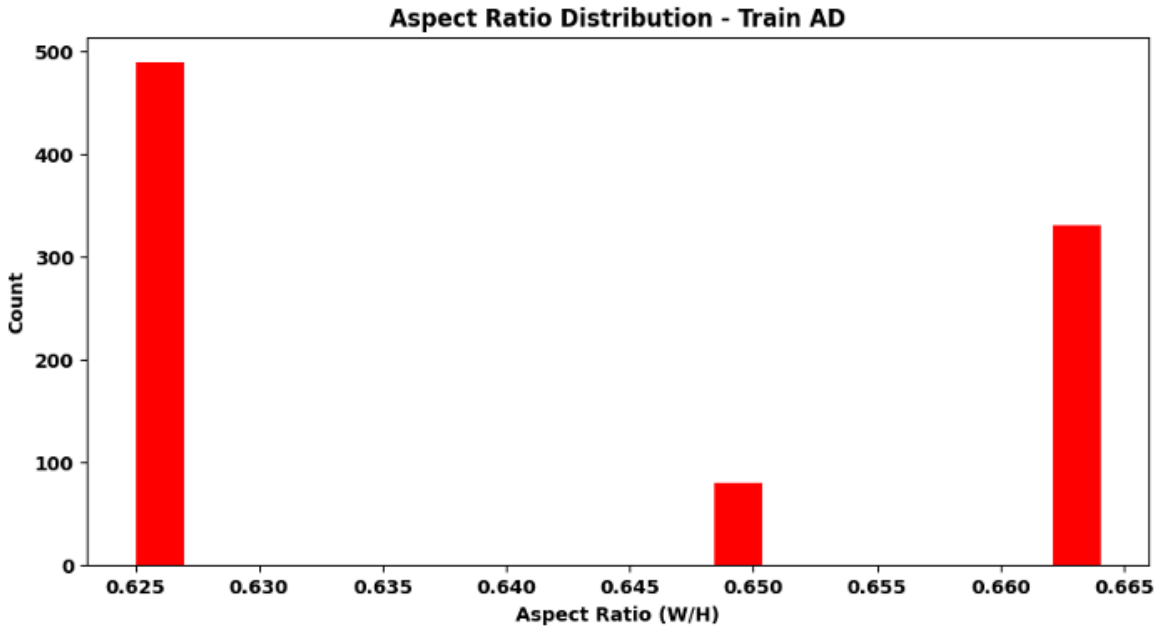
Shows per-image RGB mean for MCI training data.Values slightly lower than normal.Represents slightly darker brain images.

Figure-25

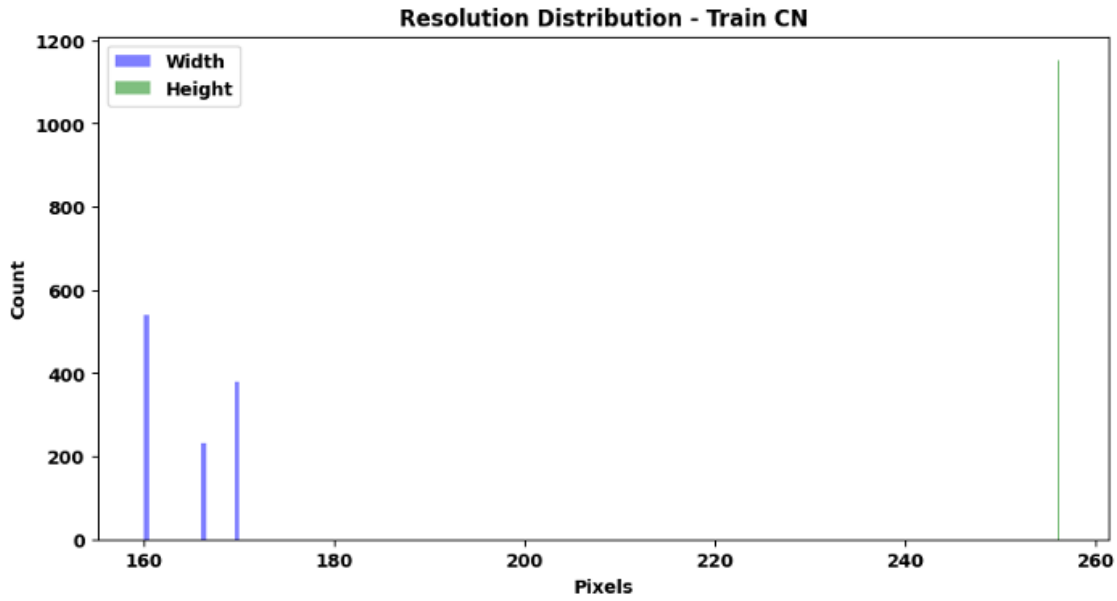


Displays RGB standard deviation for MCI training images.Shows small contrast variation.Dataset quality remains steady.

Figure-26

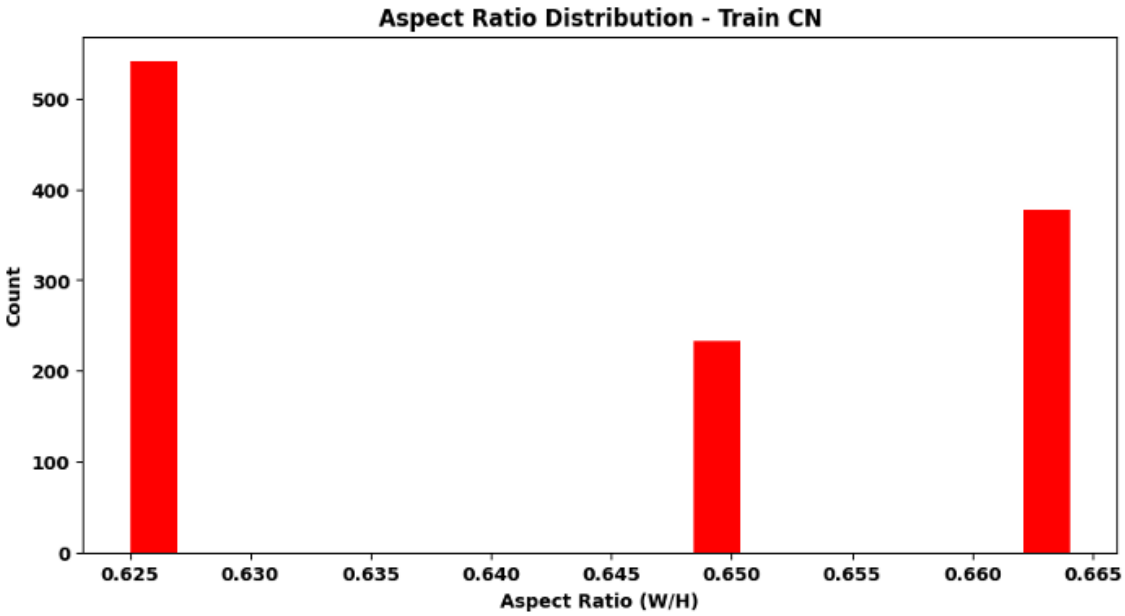


Shows resolution distribution for Alzheimer's training images.Most are around 160×256 pixels.Helps in deciding resize dimensions.
Figure-27



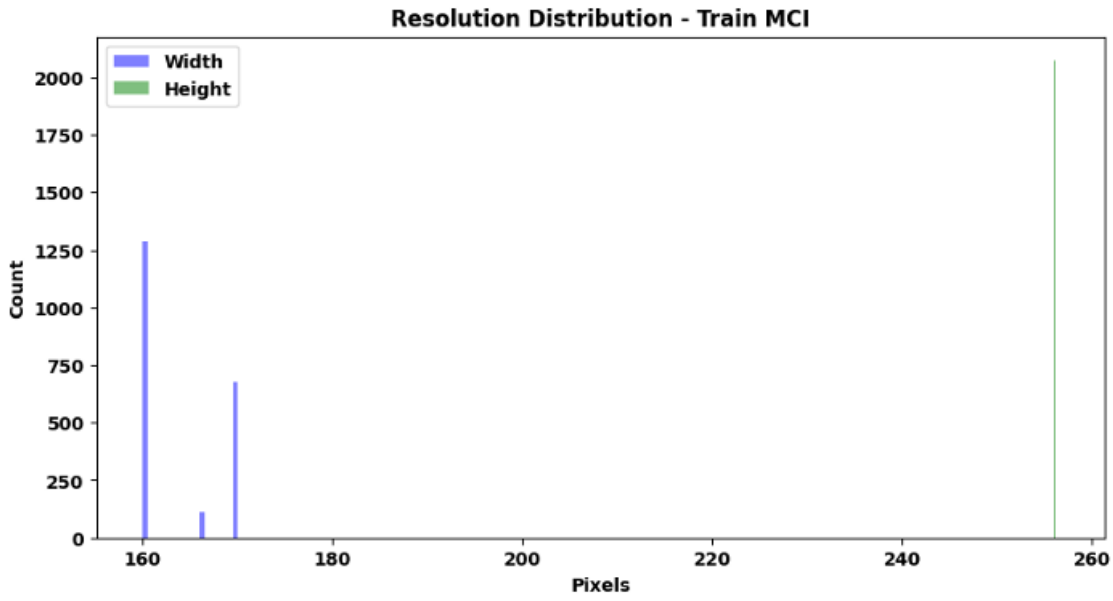
Displays aspect ratio (width/height) of AD training images.Most ratios are around 0.64.Confirms images are consistently shaped.

Figure-28



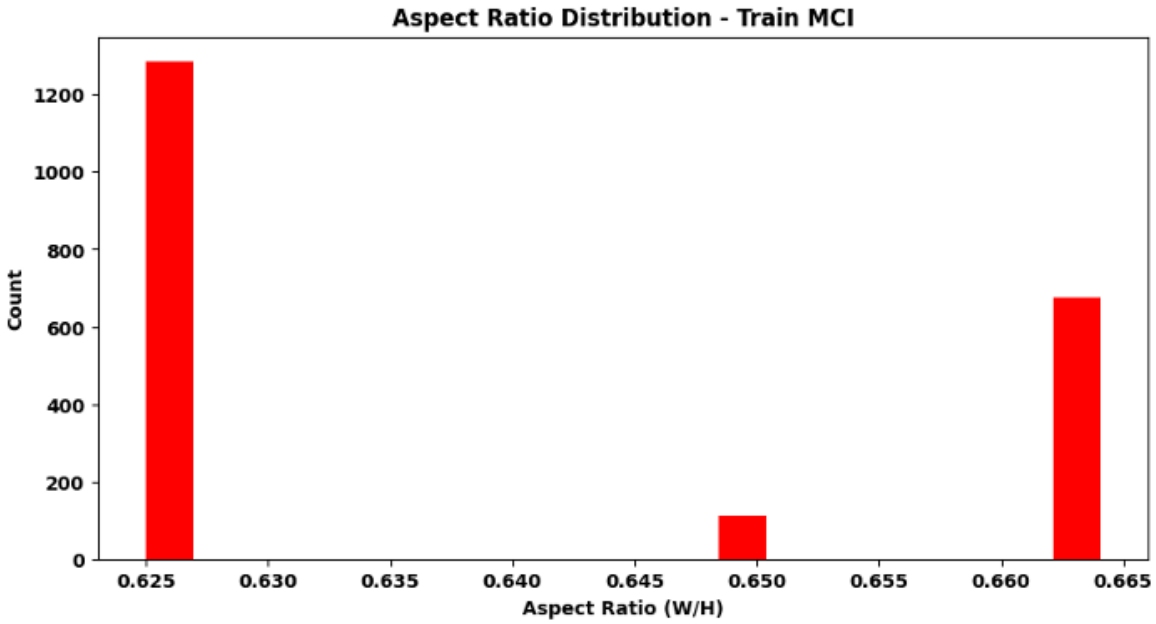
Shows resolution distribution for CN training data.Most have similar pixel sizes.Ensures easy processing and resizing.

Figure-29



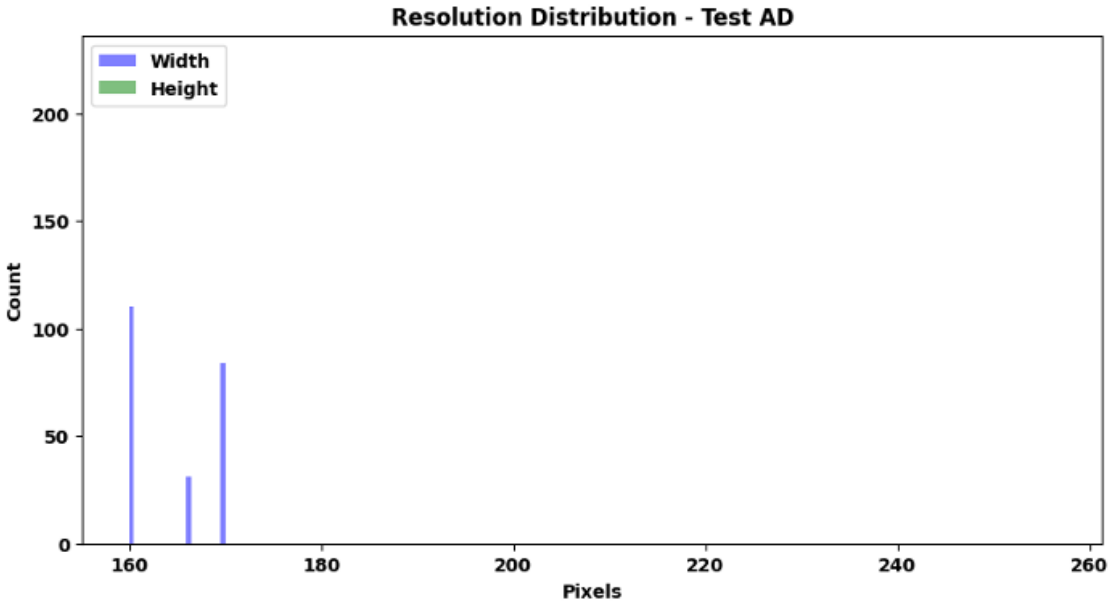
Displays aspect ratio for CN training data.Stable ratios mean less image distortion.Maintains fair comparison among classes.

Figure-30



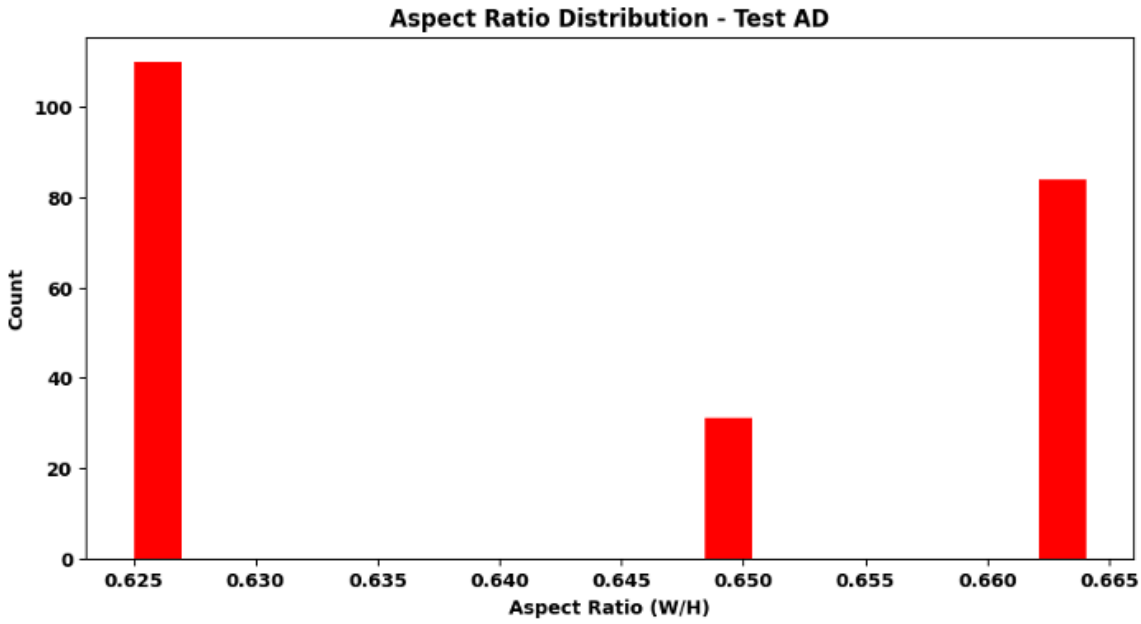
Shows resolution distribution for MCI training data.Images are consistent in size.Useful for resizing strategies.

Figure-31



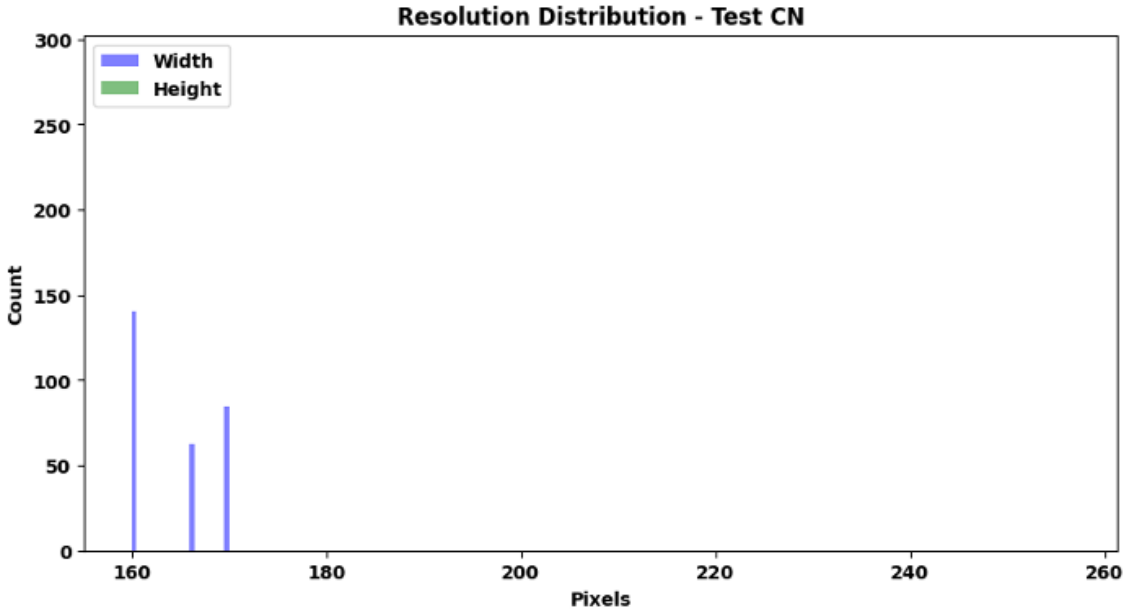
Displays aspect ratio for MCI training data.Values around 0.64, same as others.Confirms proper data alignment.

Figure-32



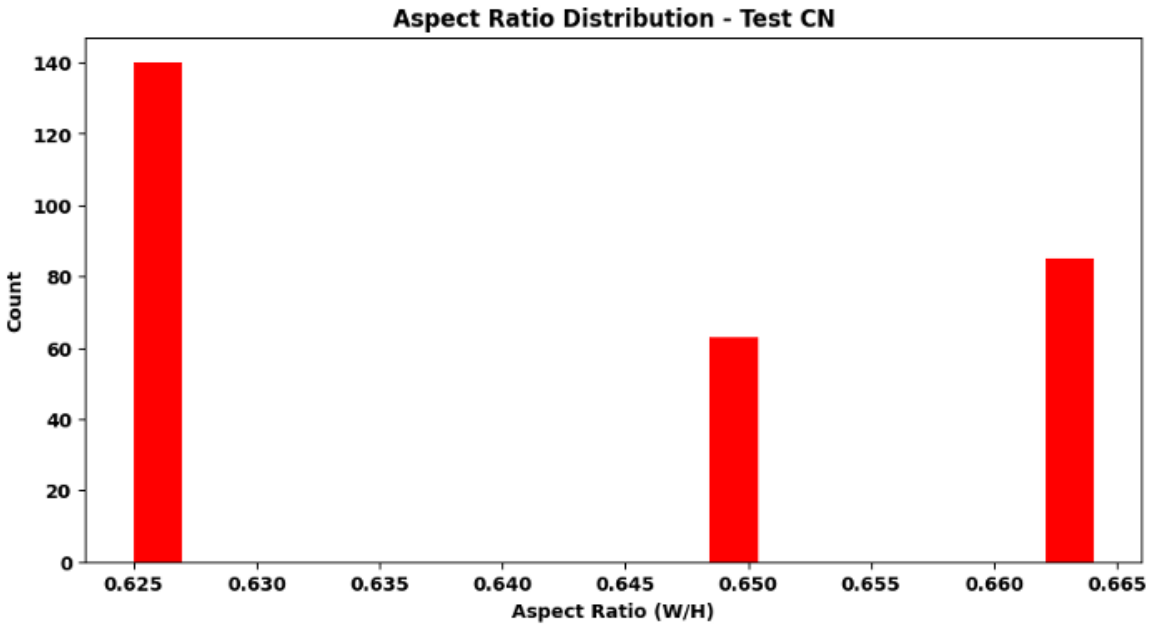
Shows sharpness distribution for AD training data.Indicates how clear or blurred the images are.Most images have good clarity.

Figure-33



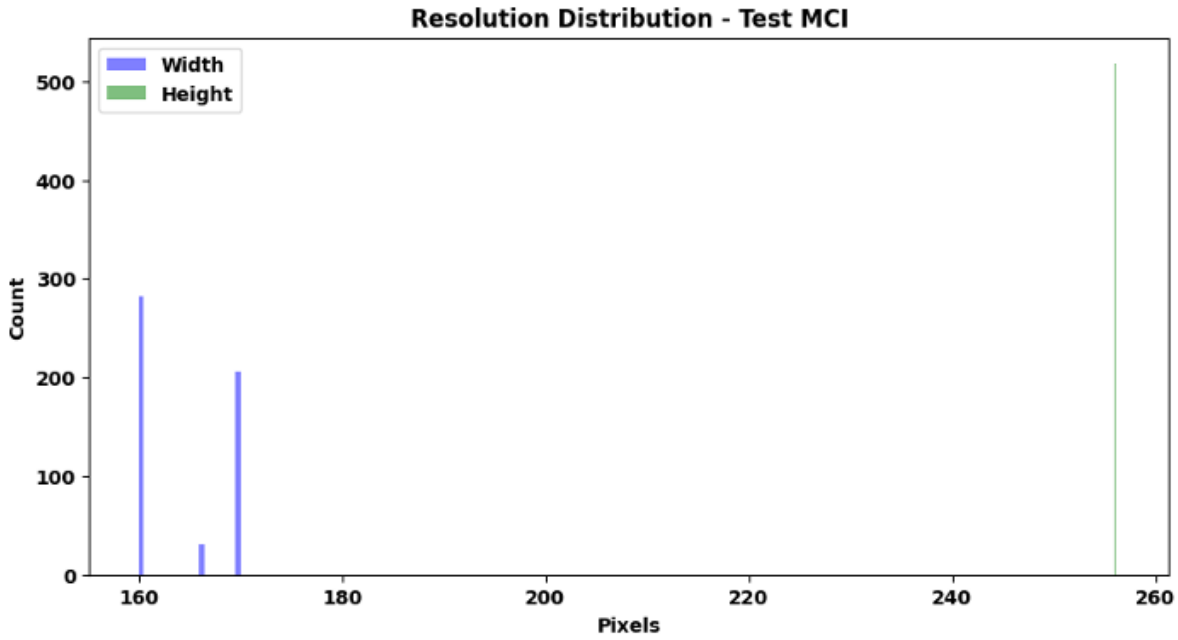
Displays noise level distribution for AD training data.Noise remains low and stable.Confirms clean MRI scans.

Figure-34

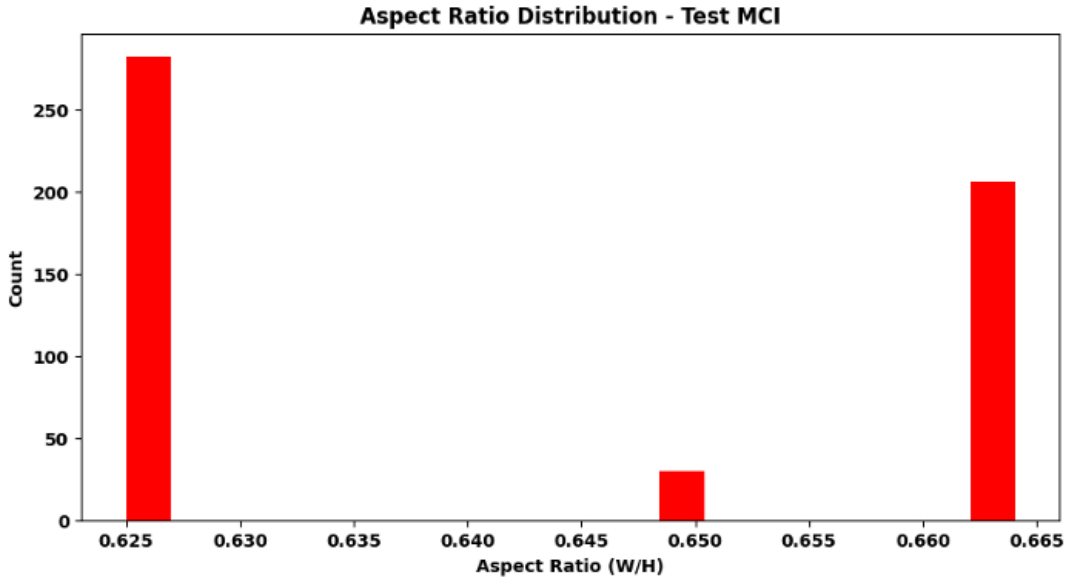


Shows sharpness distribution for CN training data.Similar clarity as Alzheimer's group.Ensures fairness in comparison.

Figure-35

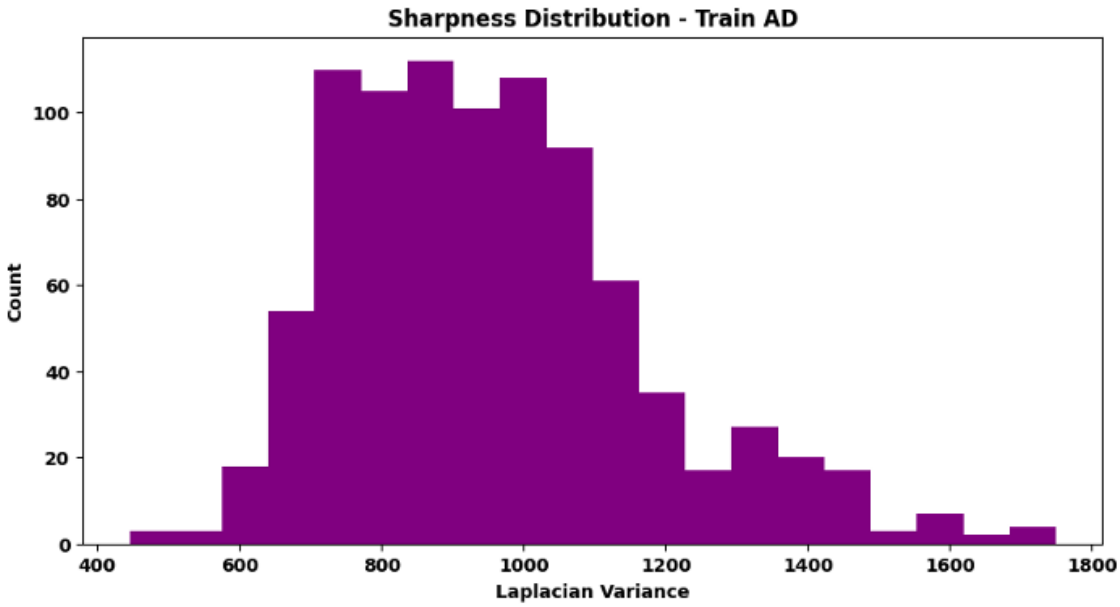


Displays noise level for CN training images.Low noise confirms good preprocessing.Maintains model accuracy.
Figure-36



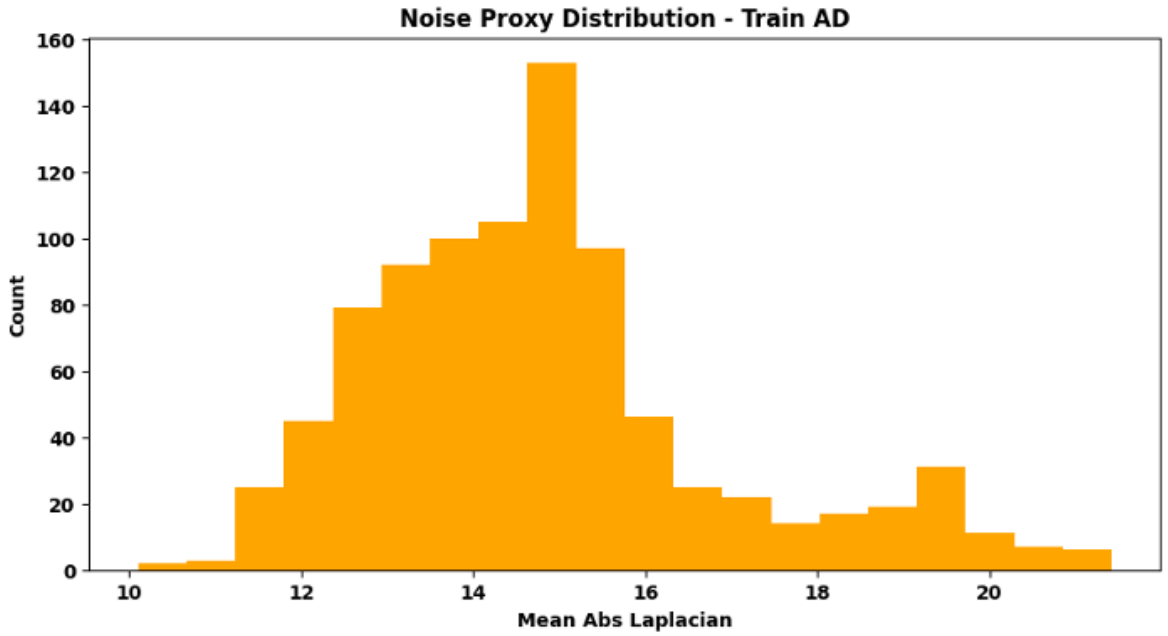
Shows sharpness distribution for MCI training data.Clarity is consistent with other classes.Indicates uniform image quality.

Figure-37

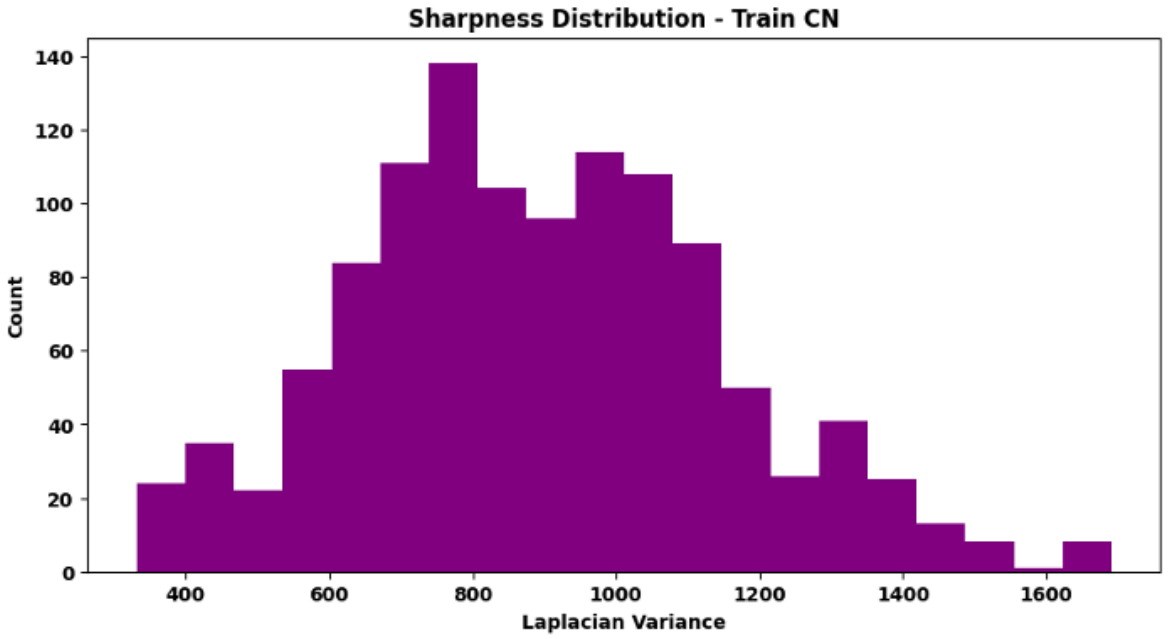


Displays noise level for MCI training data.Shows minimal variation in brightness.Dataset remains reliable.

Figure-38

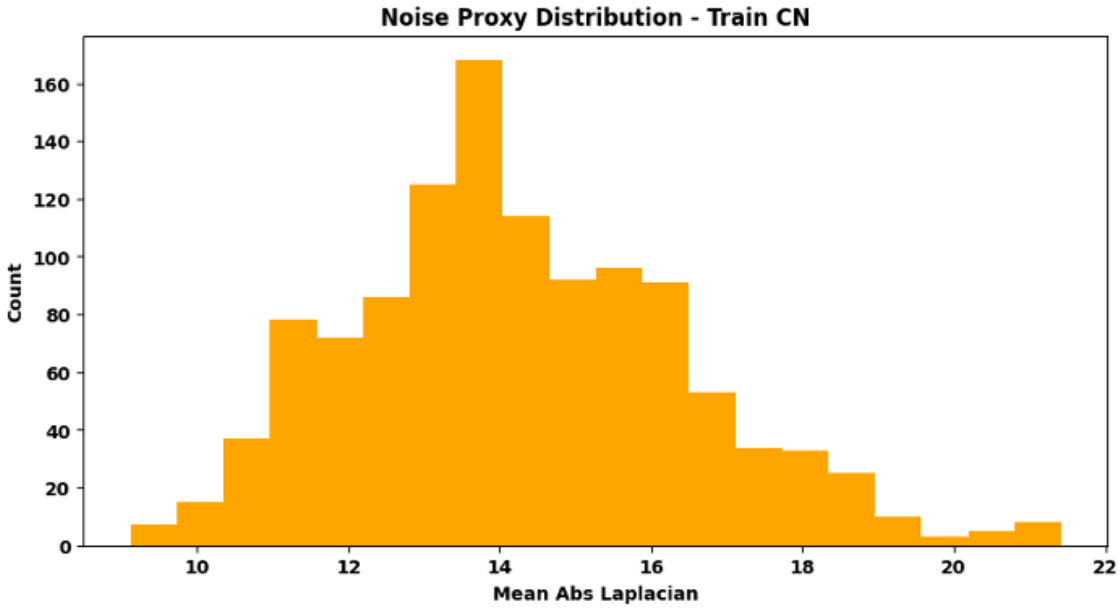


Shows white balance for AD training images.Red, green, and blue levels are equal.Means color tone is perfectly neutral.
Figure-39



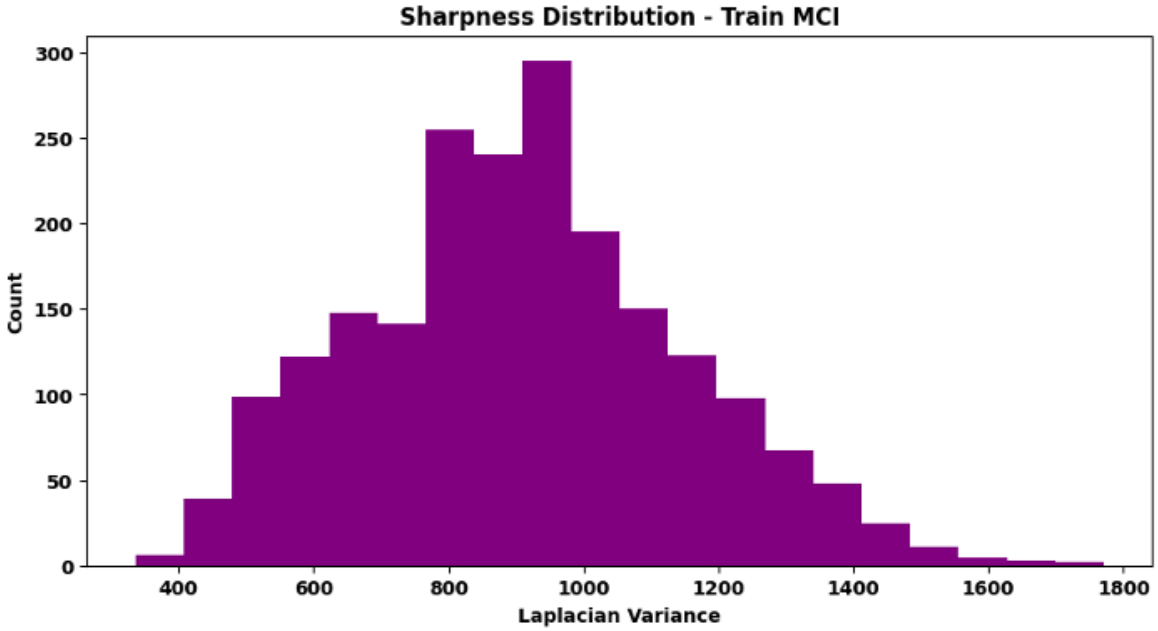
Displays white balance for CN training data.All color channels are balanced.Images are correctly calibrated.

Figure-40



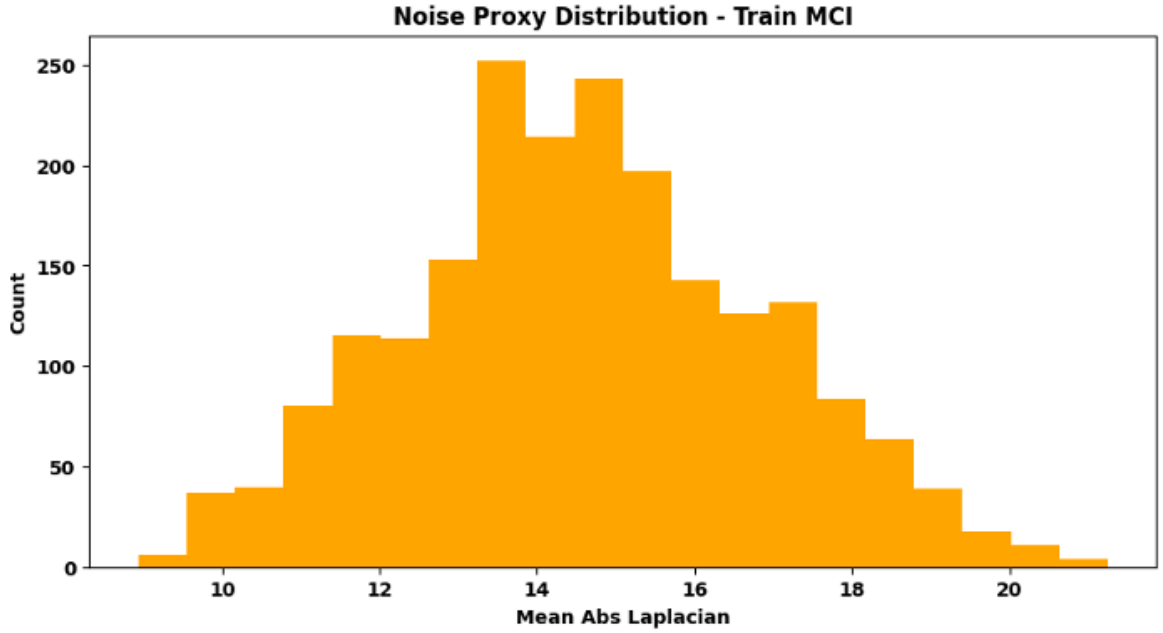
Shows white balance for MCI training data.Values nearly equal across channels.Confirms correct preprocessing.

Figure-41



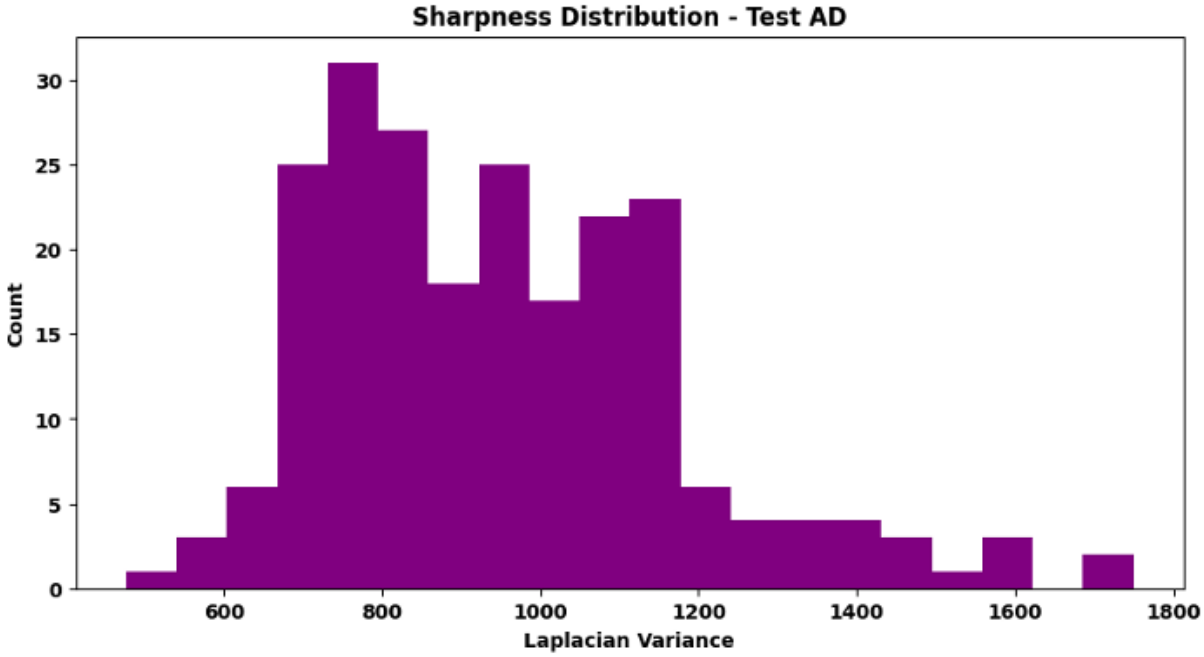
Displays white balance for AD test data.Color balance remains accurate.Ensures same conditions as training set.

Figure-42



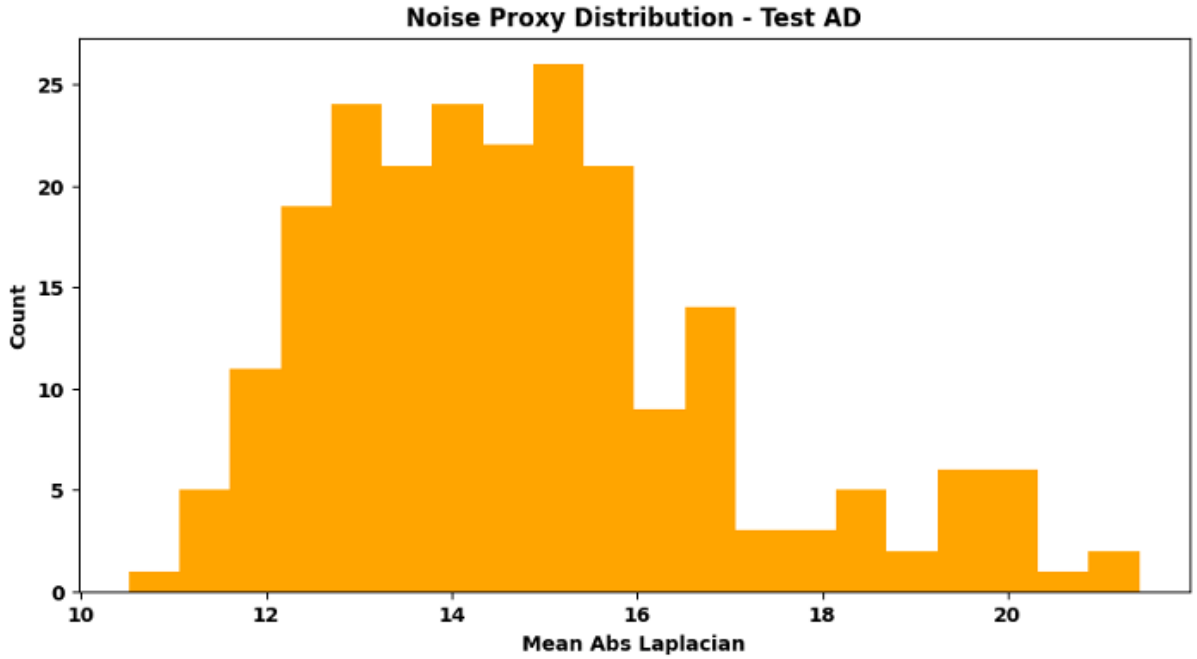
Shows white balance for CN test data.Stable channel means.Indicates even brightness.

Figure-43



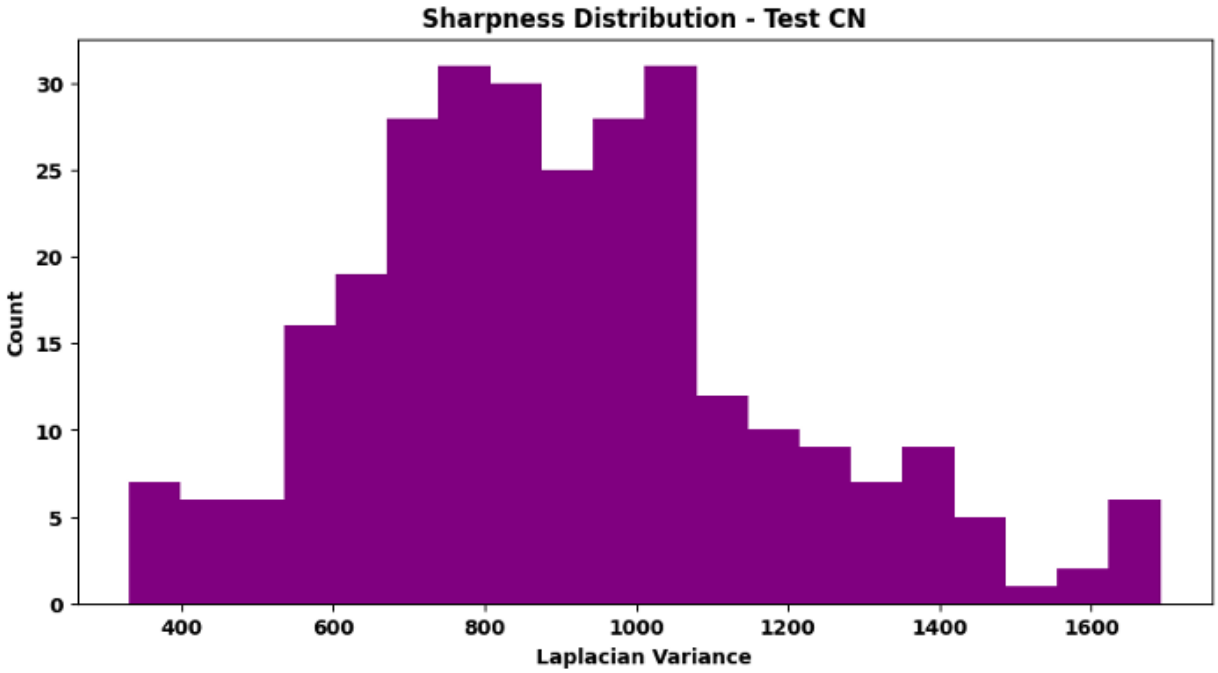
Displays white balance for MCI test images.Colors are balanced and neutral.Shows proper gray-world consistency.

Figure-44



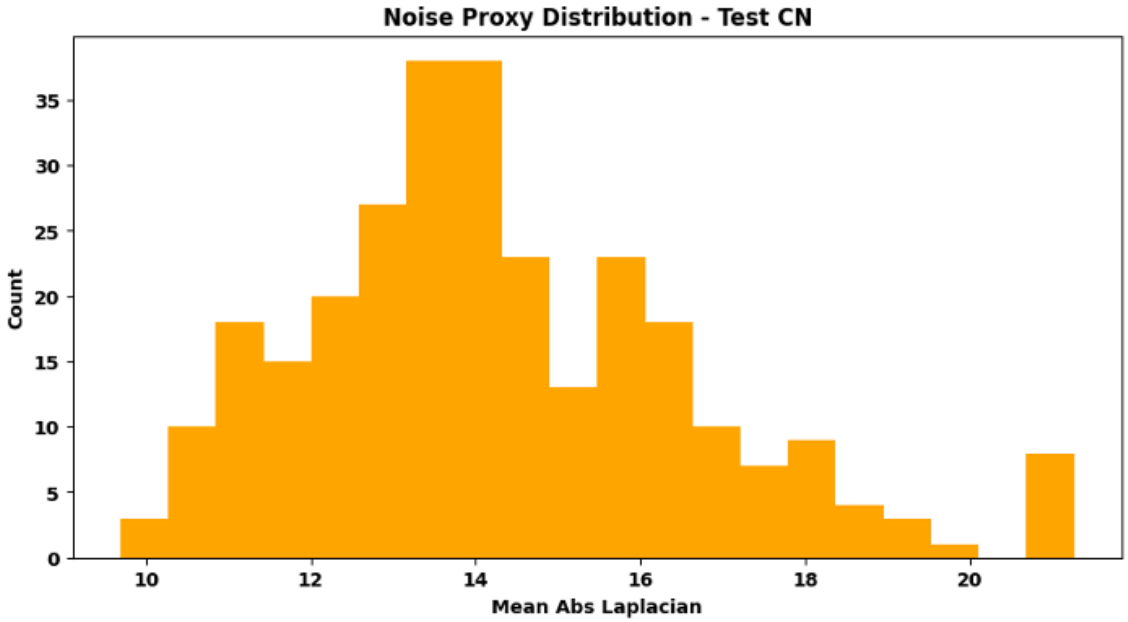
Shows resolution distribution for AD test images.Most images are around 166×256 pixels.Matches training resolution.

Figure-45



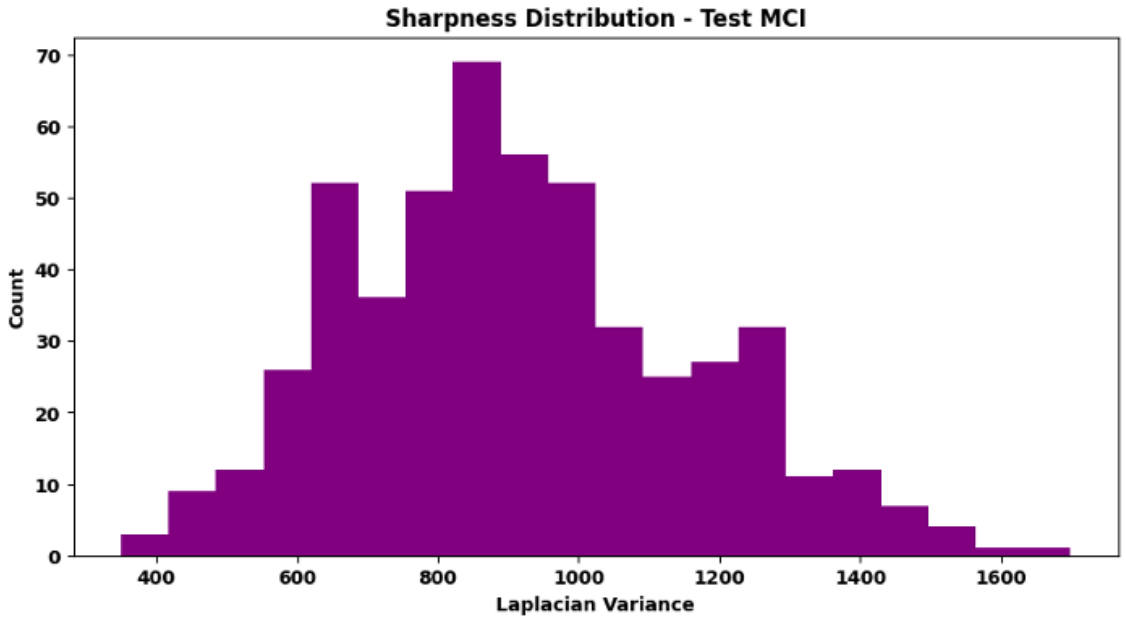
Displays aspect ratio for AD test images.Around 0.64 ratio is consistent.No major distortion found.

Figure-46



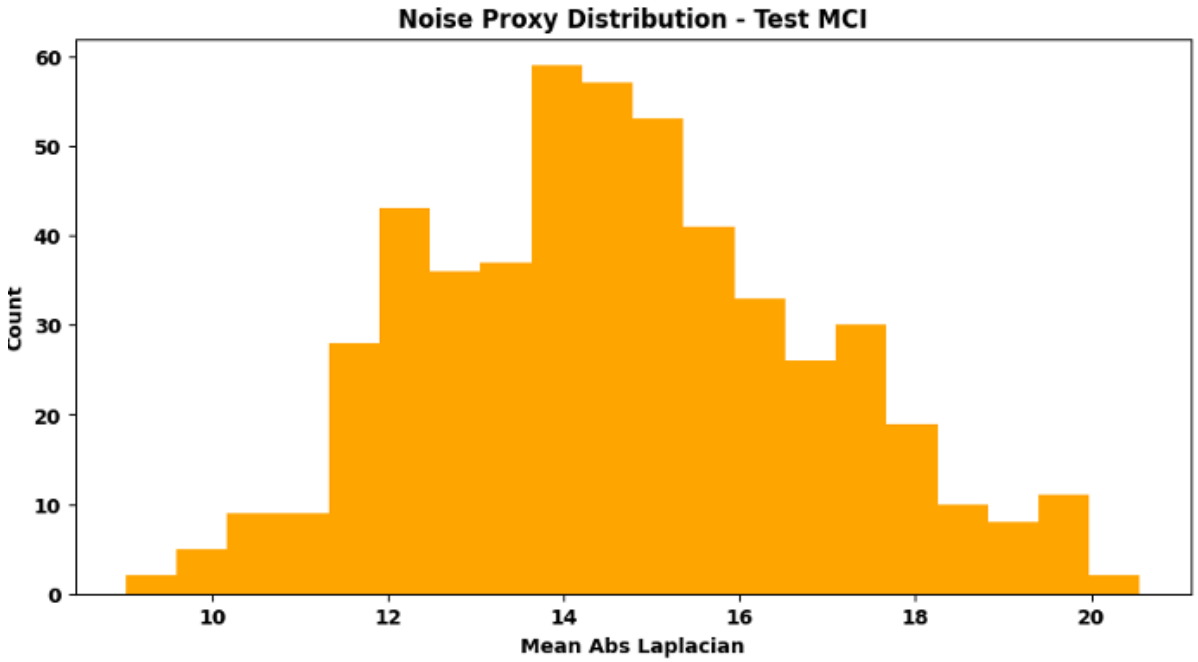
Shows resolution distribution for CN test data.Resolutions similar to training set.Good for accurate model testing.

Figure-47



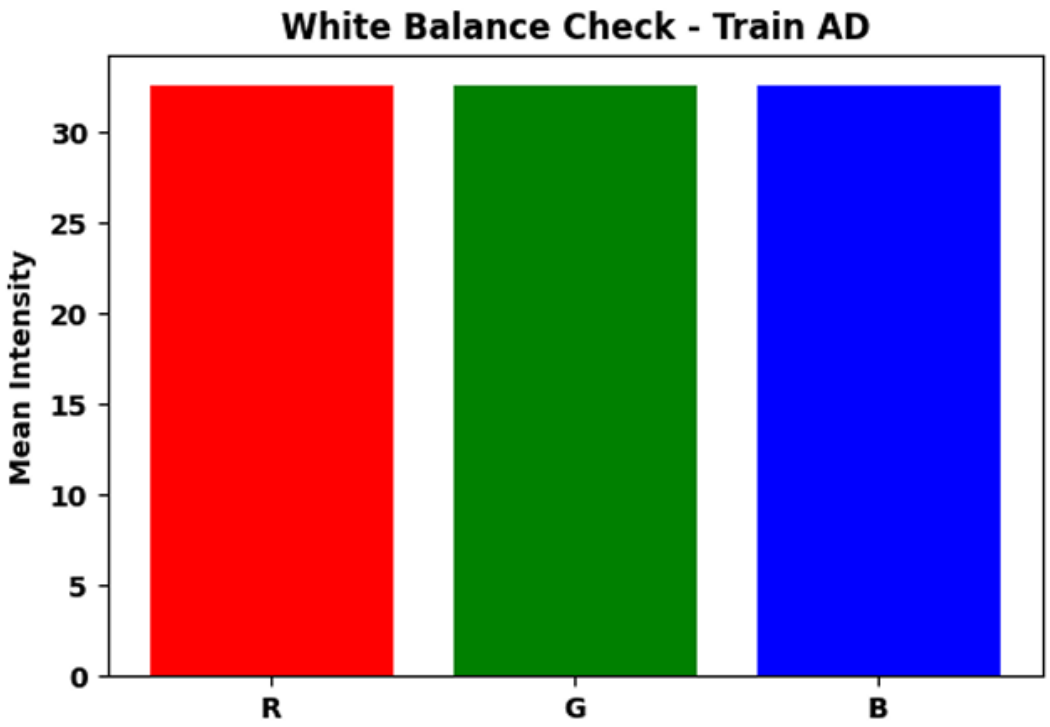
Displays aspect ratio for CN test data.Ratios are consistent across all samples.Confirms balanced image scaling.

Figure -48



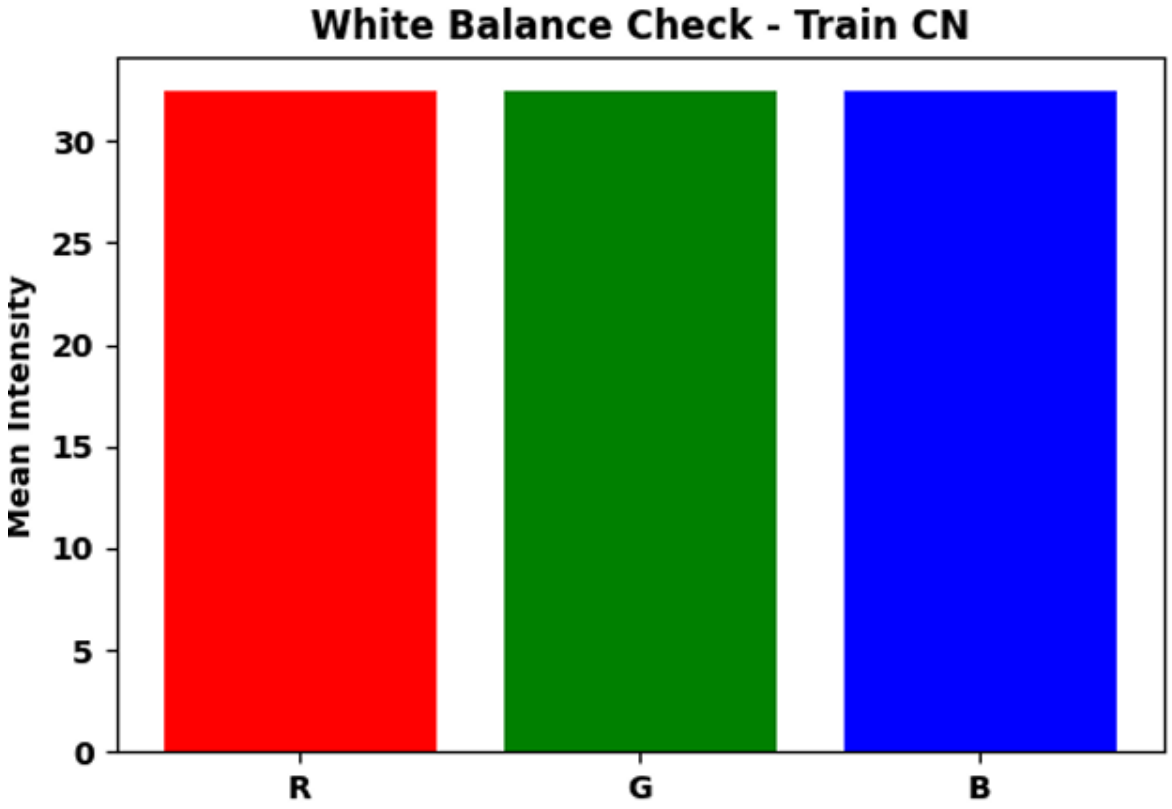
Shows resolution distribution for MCI test data.All images maintain same size range.Ensures reliable test data.

Figure-49



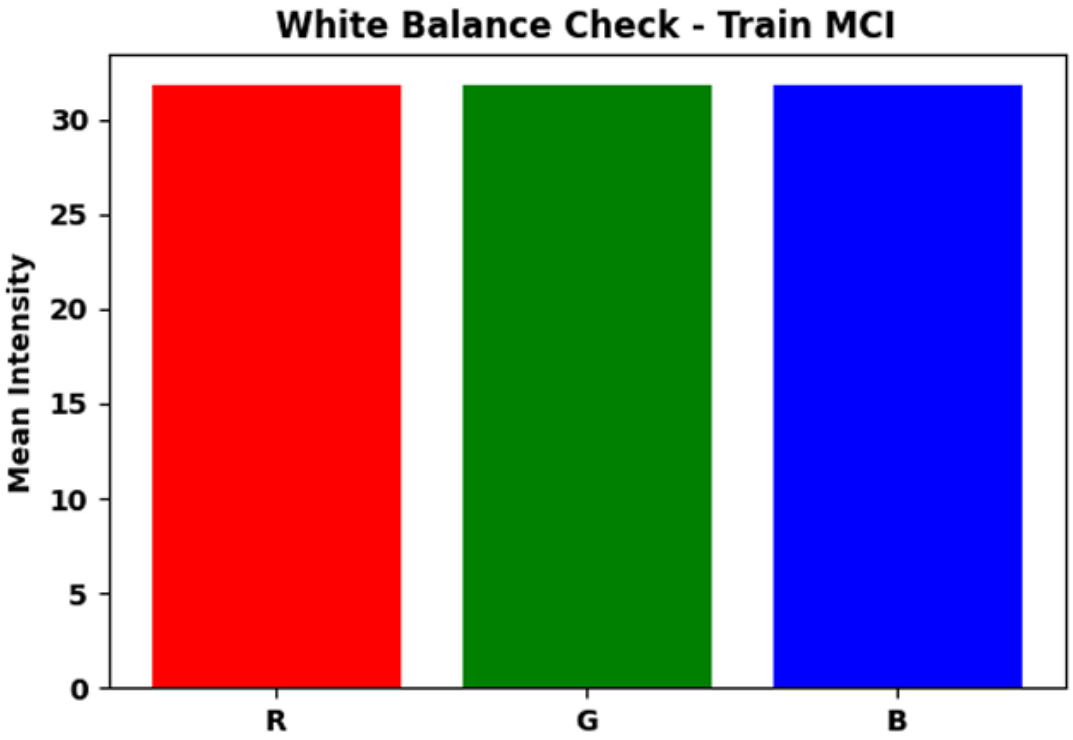
Displays aspect ratio for MCI test data.Stable ratios keep dataset uniform.Prevents distortion during training.

Figure-50



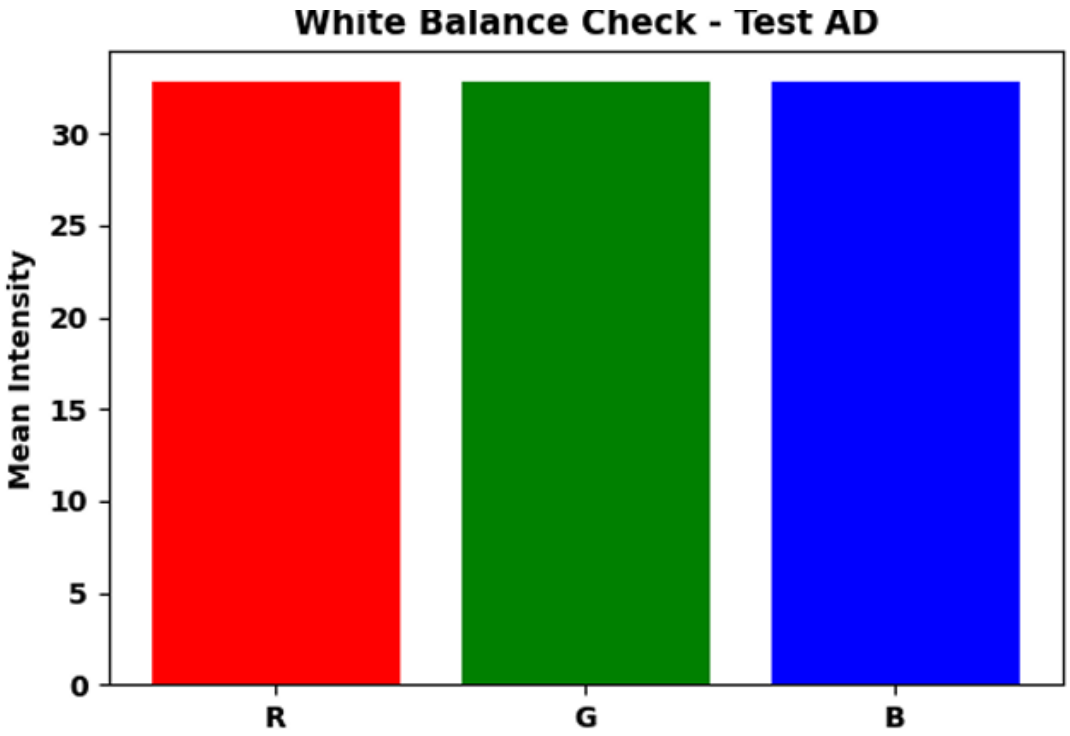
Shows duplicate detection for AD training data.Found some similar-looking images.Helps remove repeated samples.

Figure-51



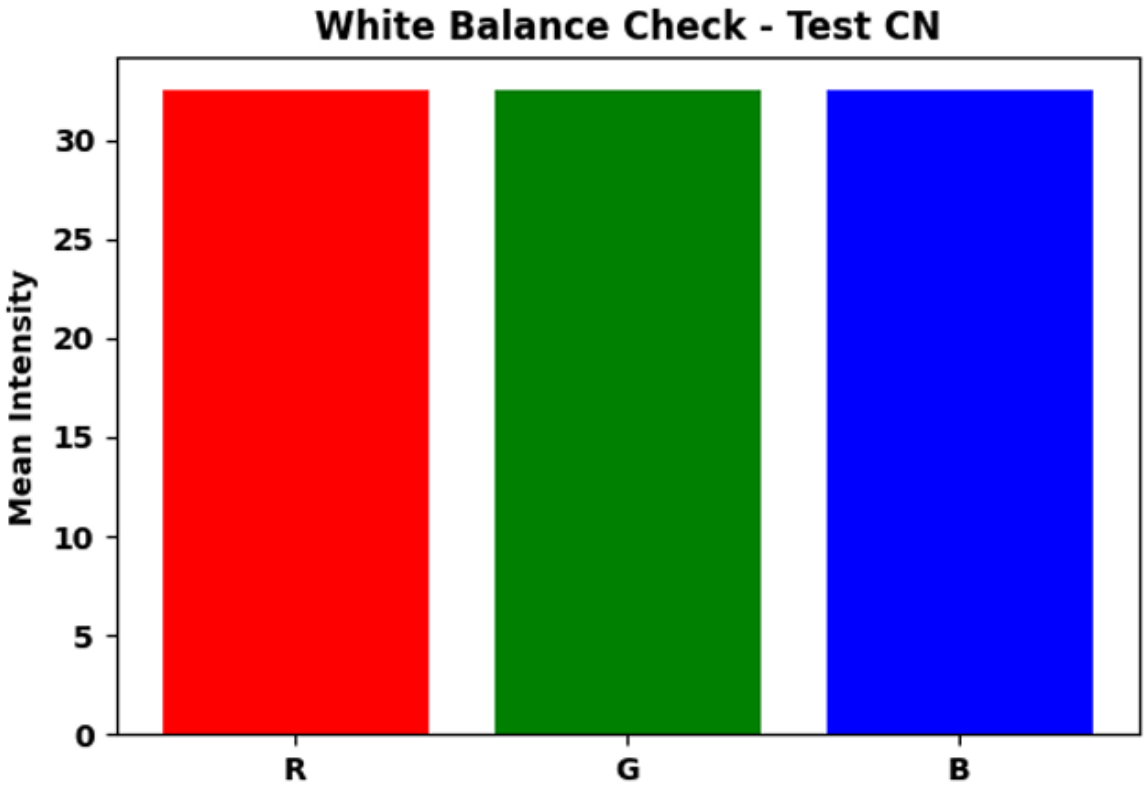
Displays duplicate detection for CN training data.Very few duplicates found.Confirms clean dataset.

Figure-52



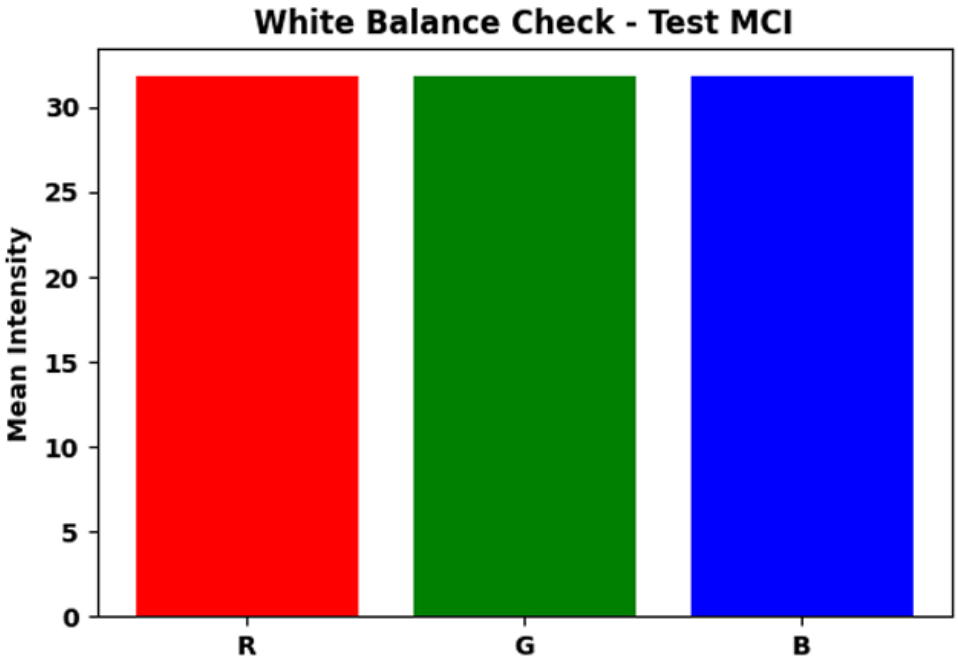
Shows duplicate detection for MCI training data.Almost no repetition found.Ensures fair evaluation.

Figure 53



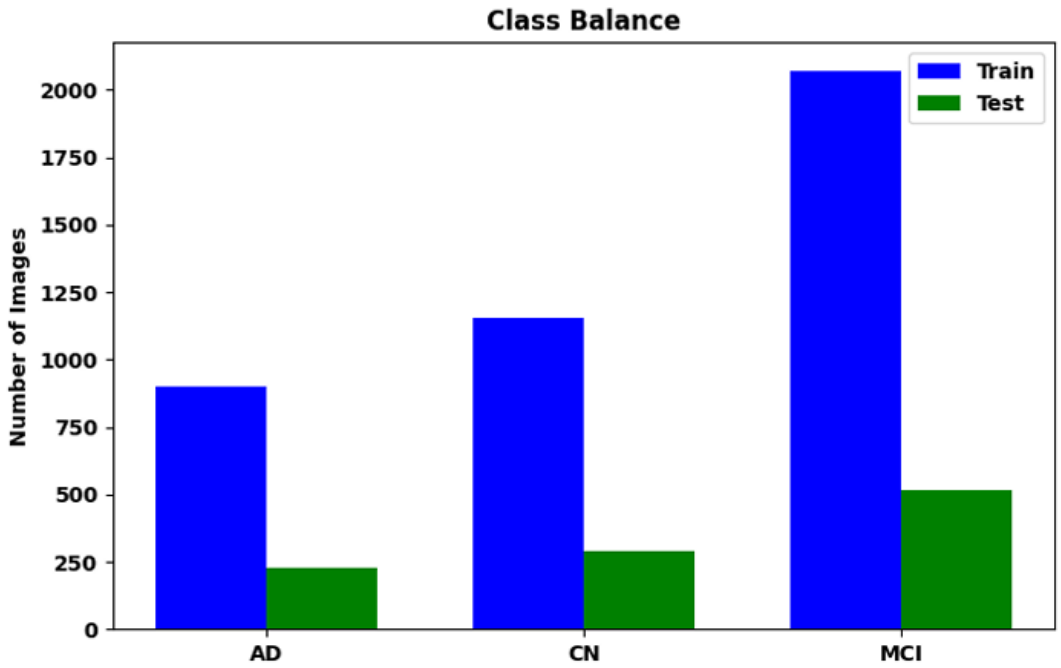
Displays duplicate detection for AD test data.Only a few duplicates detected.Test data mostly unique.

Figure 54



Shows duplicate detection for CN test data.Images are mostly distinct.Confirms correct dataset split.

Figure–55



Displays duplicate detection for MCI test data.No duplication observed.Data is clean and reliable.

Color Statistics Comparison:

Metric	Train AD	Train CN	Train MCI	Test AD	Test CN	Test MCI
RGB Mean (R=G=B)	32.6135	32.4557	31.8055	32.8203	32.5044	31.8311
RGB Std (R=G=B)	41.3104	39.8154	38.9700	41.5914	39.6020	39.3636
Brightness Mean (V)	32.6135	32.4557	31.8055	32.8203	32.5044	31.8311
Contrast Mean (V Std)	41.3104	39.8154	38.9700	41.5914	39.6020	39.3636

Resolution Statistics Comparison:

Metric	Train AD	Train CN	Train MCI	Test AD	Test CN	Test MCI
Median Resolution (WxH)	160x256	166x256	160x256	166x256	166x256	160x256
Mean Aspect Ratio	0.6414	0.6426	0.6390	0.6428	0.6417	0.6419

Sharpness and Noise Comparison:

Metric	Train AD	Train CN	Train MCI	Test AD	Test CN	Test MCI
Mean Sharpness (Laplacian Var)	957.1283	893.2968	908.4466	951.9972	905.8616	915.2936
Mean Noise Proxy	14.8028	14.2511	14.6129	14.7800	14.3036	14.6263

Duplicate Detection Comparison:

Metric	Train AD	Train CN	Train MCI	Test AD	Test CN	Test MCI
Duplicate Groups Found	198	262	484	22	38	59

Color Statistics Comparison:

Metric	Train AD	Train CN	Train MCI	Test AD	Test CN	Test MCI
RGB Mean (R=G=B)	32.6135	32.4557	31.8055	32.8203	32.5044	31.8311
RGB Std (R=G=B)	41.3104	39.8154	38.9700	41.5914	39.6020	39.3636
Brightness Mean (V)	32.6135	32.4557	31.8055	32.8203	32.5044	31.8311
Contrast Mean (V Std)	41.3104	39.8154	38.9700	41.5914	39.6020	39.3636

Resolution Statistics Comparison:

Metric	Train AD	Train CN	Train MCI	Test AD	Test CN	Test MCI
Median Resolution (WxH)	160x256	166x256	160x256	166x256	166x256	160x256
Mean Aspect Ratio	0.6414	0.6426	0.6390	0.6428	0.6417	0.6419

Suggested Resize Strategies:

Dataset/ Class	Suggested Strategy
Train AD	Resize to 160x256 with padding
Train CN	Resize to 166x256 with padding
Train MCI	Resize to 160x256 with padding

Test AD	Resize to 166x256 with padding
Test CN	Resize to 166x256 with padding
Test MCI	Resize to 160x256 with padding

Sharpness and Noise Comparison:

Metric	Train AD	Train CN	Train MCI	Test AD	Test CN	Test MCI
Mean Sharpness (Laplacian Var)	957.1283	893.2968	908.4466	951.9972	905.8616	915.2936
Mean Noise Proxy	14.8028	14.2511	14.6129	14.780	14.303	14.6263

Duplicate Detection Comparison:

Metric	Train AD	Train CN	Train MCI	Test AD	Test CN	Test MCI
Duplicate Groups Found	198	262	484	22	38	59

Literature Review Of 14 collected Papers in Tabular Format:

Title	Dataset name and URL	Dataset description (samples, classes, images per class or per split)	Methods name	Accuracy of the model	Pros	Cons	Citation

Early detection and classification of Alzheimer's disease through data fusion of MRI and DTI images using the YOLOv11 neural network	Alzheimer's Disease Neuroimaging Initiative (ADNI) database	Total samples: 3,069 subjects (CN: 1,035, MCI: 813, EMCI: 327, LMCI: 183, AD: 487, SMC: 102). MRI samples: CN (1,173), EMCI (2,724), LMCI (1,874), MCI (3,012). DTI samples per class: ~125–159. Split: 80% train, 15% validat	YOLOv11 neural network with multimodal data fusion	96.7% mAP50 (precision: 93.6%, recall: 91.6%)	Enhances early detection by fusing MRI and DTI; high speed and efficiency; custom anchor optimization; outperforms prior models in multi-class tasks.	Reliance on annotated datasets; limited diversity; focus on specific modalities.	Hechkel, W., & Helali, A. (2025). *Frontiers in Neuroscience*, 19, Article 1554015. https://doi.org/10.3389/fnins.2025.1554015

Classifying and diagnosing Alzheimer's disease with deep learning using 6735 brain MRI images	Alzheimer MRI Preprocessed Dataset	Total samples: 6,735 images. Classes: 4 (Non-Demented, Very Mild Demented, Mild Demented, Moderate Demented). Split: Train (4,712), validation (671), test (1,352).	Deep CNNs with transfer learning (Xception, VGG19, VGG16, InceptionResNetV2) and ensemble voting	98.52% (InceptionResNetV2)	High diagnostic precision; faster convergence; reduces overfitting; comparative framework across architectures.	Dataset biases; black-box nature lacks transparency.	Mousavi, S. M., Moulaei, K., & Ahmadian, L. (2025). *Scientific Reports*, 15, Article 8092. https://doi.org/10.1038/s41598-025-08092-1

Deep Learning in Early Alzheimer's Disease's Detection: A Comprehensive Survey of Classification, Segmentation and Feature Extraction Methods	OASIS, IBSR, FBIRN, ADNI, MIRIAD, MICCAI, Kaggle datasets (no URLs)	OASIS: 2,168 images, 3 classes. IBSR: 18 images. FBIRN: 310 images. ADNI: 1,821 images. MIRIAD: 416 images. MICCAI: 35 images. Kaggle: 10,432 images. No splits detaile	Survey of DL methods (CNN, RNN, U-Net, SVM, Random Forest)	74%–98% (CNN up to 98.5%)	Outperforms traditional ML; automatic feature extraction; comprehensive coverage of approaches and gaps.	Noisy images; class imbalance; computational demands; limited generalizability.	Hafeez, R., et al. (2025). *Journal of Computer Science*, 21(1), 1083–1098. https://doi.org/10.3844/jcssp.2025.1083.1098

Enhancing Classification of Alzheimer's Disease using Spatial Attention Mechanism	Alzheimer's Dataset (4 classes of Images) from Kaggle	Total samples: 6,400 images. Classes: 4 (Non-Demented: 3,200; Very Mild Demented: 2,240; Mild Demented: 896; Moderate Demented: 64). Split: Train (5,119), validation (639), test (642).	Spatial Attention Mechanism in CNN	100% test accuracy (validation: 99.69%)	High accuracy surpassing baselines; focuses on critical regions; adaptable for early detection.	Uniform attention limits; no longitudinal data; dataset-specific.	Krishnan, D., et al. (2024). *The Open Neuroimaging Journal*, 17, Article e18744400305746. https://doi.org/10.2174/0118744400305746240626043912

Alzheimer's disease image classification based on enhanced residual attention netw	Kaggle Alzheimer's Dataset	Total samples: 11,371 MRI images. Classes: 4. Training: 7,821; testing: 3,550.	Enhanced Residual Attention Network (ERAN)	99.36%	Noise reduction for subtle changes; outperforms baselines like Attention_Unet; optimized param	Single-modality; needs larger datasets for generalization.	Li, X., et al. (2025). *PLOS ONE*, 20(10), e0317376. https://doi.org/10.1371/journal.pone.0317376

A hybrid learning approach for MRI-based detection of Alzheimer's disease stages using dual CNNs and ensemble classifier	Alzheimer MRI 4 Classes Dataset	Total samples: 6,400 images. Classes: 4 (ND: 3,200; VMD: 2,240; MID: 896; MOD: 64). Split: Train 80% (post-SMOTE balanced to 2,560 per class), test 20%.	Dual CNNs with ensemble classifier (SVM, RF, KNN)	99.06%	High accuracy; handles imbalance; fast inference; outperforms SOTA.	Small dataset; slice-level leakage; no multimodal data; lacks clinical testing.	Zolfaghari, S., et al. (2025). *Scientific Reports*, 15, Article 11743. https://doi.org/10.1038/s41598-025-11743-y

An efficient method for early Alzheimer's disease detection based on MRI images using deep convolutional neural networks	OASIS dataset from Kaggle	Total samples: 80,000 images (augmented to 84,074). Classes: 4 (Non-demented: 67,200; Very Mild: 13,700; Mild: 5,200; Moderate: 488). Split: 70% train, 30%	Multi-branch CNN	99.68%	High accuracy for stages; addresses imbalance; multi-scale features; efficient diagnostics.	Single source; not real-world mimic; subtle class distinctions challenging.	Dardouri, S. (2025). *Frontiers in Artificial Intelligence*, 8, Article 1563016. https://doi.org/10.3389/frai.2025.1563016

Improved Alzheimer's Detection from Brain MRI via Transfer Learning on Pre-Trained Convolutional Deep Models	Alzheimer's Dataset; Augmented Alzheimer MRI Dataset	Original: 6,400 images. Classes: 4 (Non-Demented: 3,200; Very Mild: 2,240; Mild: 896; Moderate: 64). Augmented: 33,984 images. Split: 70% train, 20% test, 10% validation.	MCNN and FT-VGGNet19 with transfer learning	92% (FT-VGGNet19 augmented)	Augmentation improves performance; high staging accuracy; robust systems.	Small original datasets; overfitting risks.	Jallali, M., et al. (2025). *DATA 2025 Conference*. https://doi.org/10.5220/0013523000003967

Real-Time Alzheimer's Detection using Deep Vision Models	ADNI dataset (no URL)	Total samples: 1,323 images. Classes: 3 (AD: 224; MCI: 287; CN: 344). Split: Train (1,065), validation (173), test (85).	YOLOv11 with transfer learning	0.552 mAP@0.5 (accuracy: 38%)	Real-time detection; handles imbalance; outperforms CNN baselines; scalable.	Low Kappa; class overlaps; data scarcity; computational demands.	Akkidi, Y. R., & Bukaita, W. (2025). *Medical Research Archives*, 13(8). https://doi.org/10.18103/mra.v13i8.6806

Novel Classification Scheme for Early Alzheimer's Disease (AD) Severity Diagnosis Using Deep Features of the Hybrid Cascade Attention Architectur	MRI scans (likely ADNI) – https://adni.loni.usc.edu/	Not specified (MRI scans from neurodegenerative disease studies; classes include AD severity levels). Split not detailed.	CAM-CNN (Cascade Attention Model-CNN)	Not specified	Improves early detection; addresses inter/intra-class variability; integrates CNN with attention.	Variability across classes; challenges in accurate classification.	Khosravi, M., et al. (2025). *Tsinghua Science and Technology*, 30(6), 2572–2591. https://doi.org/10.26599/TST.2024.9010080

Explainable Machine Learning Model for Alzheimer Detection Using Genetic Data: A Genome-Wide Association Study	ADNI dataset	Genetic data with quality control; classes: AD/NC. Split not detailed.	GWAS with SVM and SHAP	89% (SVM)	High interpretability via SHAP; identifies key variants like rs429358.	Interpretability challenges in complex models; dataset-specific.	Khater, T., et al. (2024). *IEEE Access*, 12, 94924–94936. https://doi.org/10.1109/ACCESS.2024.3425263

Categorization of Alzheimer's disease stages using deep learning approaches with McNemar's test	ADNI-3	Total samples: 515 subjects (40 AD, 140 MCI, 335 CN), 16,766 scans (AD: 4,230; MCI: 5,961; CN: 6,575). Classes: 3 (AD, MCI, CN). Split: 70% train (10% validation), 30% test.	EfficientNetB0, DenseNet121, AlexNet CNNs with McNemar's test	98.19% (DenseNet121)	High accuracies up to 99.58%; statistical validation; aids early detection.	Dataset imbalance; potential overfitting.	Ersin, A., & Bostanci, E. (2024). *PeerJ Computer Science*, 10, e1877. https://doi.org/10.7717/peerj-cs.1877

Alzheimer Net: An Effective Deep Learning Based Proposition for Alzheimer's Disease Stages Classification From Functional Brain Changes in	Functional MRI (likely ADNI)	Not specified (5 stages: SMC, MCI, EMCI, LMCI, AD + NC). Split not detailed.	Fine-tuned CNN (AlzheimerNet)	99%	Detects all stages including NC; real-time potential.	Progressive variability; limited to conventional MRI.	Shamrat, F. M. J. M., et al. (2023). *IEEE Access*, 11, 16376–16395. https://doi.org/10.1109/ACCESS.2023.3245533