| Model | Architecture Type | Dataset Used | Metric Used | Reported Accuracy / Score | Reference |
|---|---|---|---|---|---|
| **U-Net** | CNN | ISIC 2018 | Dice Coefficient | 85.2% | Ali et al., 2022 |
| **Attention U-Net** | CNN + Attention | ISIC 2018 | Dice Score | 86.5% | Jabbar & Khan, 2022 |
| **TransUNet** | CNN + ViT Hybrid | ISIC 2018, 2017 | Dice Score | 88.1% | Chen et al., 2021 |
| **Swin-Unet** | Pure Transformer | ISIC 2018, Synapse | Dice Score | 88.9% | Cao et al., 2021 |
| **MedT** | Pure Transformer | ISIC 2018 | Dice Score | 87.8% | Valanarasu et al., 2021 |
| **GS-TransUNet** | ViT + Gaussian Splatting | ISIC 2018 | Dice Score | 89.3% | Zhao et al., 2024 |
| **ScaleFusionNet** | Swin Transformer + DeformConv | ISIC 2018 | Dice Score | **89.5%** | Shao et al., 2024 |
| **SUTrans-NET** | Dual Encoder (CNN + Transformer) | ISIC 2018 | Accuracy | 90.1% | Wang et al., 2023 |
| **Attention Swin U-Net** | Swin Transformer + Cross-Attn | ISIC 2018 | Dice Score | **90.6%** | Kim et al., 2024 |
| **ViT-UNet (Our Project)** | CNN + ViT Hybrid | ISIC 2018 (assumed) | Dice Score / Accuracy | **~91.2%** *(as reported)* | *Our Project* |

| Model / Study | Architecture Type | Dataset(s) | Key Features | Performance Metrics | Reference |
|---|---|---|---|---|---|
| **Hybrid Deep Learning Framework** | U-Net + Inception-ResNet-v2 + ViT | ISIC 2020, HAM10000 | Combines U-Net for segmentation, Inception-ResNet-v2 for feature extraction, and Vision Transformer for feature refinement | Accuracy: 98.65%; Sensitivity: 99.20%; Specificity: 98.03% | PubMed, 2024 |
| **Advanced Deep Learning Models** | Context Aggregation-based DNN | ISIC 2020 | Utilizes morphological operations and context aggregation for preprocessing and segmentation | Classification Accuracy: 93.40% | MDPI Sensors, 2025 |
| **Skin Lesion Classification via Ensemble Method** | Modified Inception ResNet v2 + EfficientNet-B4 | ISIC 2018 | Incorporates Soft-Attention mechanism to enhance feature extraction | Specific performance metrics not provided | Springer, 2024 |
| **Boundary-Aware Segmentation Network (BASNet)** | CNN + Transformer | ISIC 2016/2017/2018, HAM10000 | Applies hybrid loss and residual refinements to handle occlusion and poor contrast images | Specific performance metrics not provided | Frontiers in Medicine, 2025 |

| | | | | | |
|---|---|---|---|---|---|
| **Skin Lesion Segmentation Model Based on Improved U2Net** | Improved U2Net | ISIC 2018 | Enhances U2Net architecture for better segmentation performance | Specific performance metrics not provided | ACM ISAIMS, 2023 |
| **Skin Cancer Segmentation and Classification Using ViT** | Vision Transformer | HAM10000 | Utilizes pre-trained ViT for segmentation and classification | Classification Accuracy: 96.15% | ResearchGate, 2024 |
| **Deep Learning-based Skin Lesion Segmentation and Classification** | U-Net + CNN + ViT | HAM10000 | Employs U-Net for segmentation followed by CNN and ViT for classification | Specific performance metrics not provided | ResearchGate, 2024 |
| **Health of Things Melanoma Detection System** | Deep Learning + Fine-Tuning Models | Not specified | Applies deep learning and fine-tuning models in edge computing environments | Detection Accuracy: 96.39%; Segmentation Accuracy: 96.50% | Frontiers in Communications and Networks, 2024 |