**S1. Deep Learning Models**

**S1.1 Object Detection**

**S1.1.1 YOLOV5**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Figure S1.** Confusion Matrix of YOLOV5

**A graph of a graph of a function

AI-generated content may be incorrect.**

**Figure S2.** Precision, Recall, Accuracy & loss Graphs

**Table S1.** Classification Report of YOLOV5

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Class** | **Images** | **Instances** | **Precision (P)** | **Recall (R)** | **mAP@0.5** | **mAP@0.5:0.95** |
| **all** | 284 | 8369 | 0.855 | 0.774 | 0.832 | 0.556 |
| **Adaptor** | 284 | 48 | 0.819 | 0.752 | 0.82 | 0.564 |
| **Allen Key** | 284 | 6 | 0.856 | 0.989 | 0.972 | 0.723 |
| **Bit Holder** | 284 | 8 | 0.602 | 0.125 | 0.259 | 0.189 |
| **Bit Sockets** | 284 | 601 | 0.979 | 0.955 | 0.97 | 0.697 |
| **Bits** | 284 | 534 | 0.929 | 0.513 | 0.754 | 0.354 |
| **Deep Socket** | 284 | 1981 | 0.889 | 0.766 | 0.805 | 0.605 |
| **Extension Bar** | 284 | 500 | 0.951 | 0.868 | 0.932 | 0.605 |
| **Flex Handle** | 284 | 13 | 0.734 | 0.846 | 0.732 | 0.456 |
| **Rachet-Handels** | 284 | 200 | 0.954 | 0.87 | 0.913 | 0.652 |
| **S Handle** | 284 | 4 | 0.502 | 0.5 | 0.711 | 0.386 |
| **Sockets** | 284 | 3961 | 0.96 | 0.87 | 0.915 | 0.647 |
| **Spanners** | 284 | 12 | 0.938 | 1.0 | 0.995 | 0.647 |
| **long bit socket** | 284 | 320 | 0.898 | 0.869 | 0.925 | 0.57 |
| **universal joint** | 284 | 181 | 0.959 | 0.908 | 0.94 | 0.688 |

**S1.1.2 YOLOV8**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Figure S3.** Confusion Matrix of YOLOV8

**A group of graphs showing the growth of a number of data

AI-generated content may be incorrect.**

**Figure S4.** Precision, Recall, Accuracy & loss Graphs

**Table S2.** Classification Report of YOLOV8

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Class | Images | Instances | Precision (P) | Recall (R) | mAP@0.5 | mAP@0.5:0.95 |
| all | 284 | 8369 | 0.913 | 0.896 | 0.926 | 0.731 |
| Adaptor | 34 | 48 | 0.927 | 0.791 | 0.853 | 0.707 |
| Allen Key | 6 | 6 | 0.83 | 1.0 | 0.995 | 0.873 |
| Bit Holder | 7 | 8 | 0.896 | 0.75 | 0.758 | 0.609 |
| Bit Sockets | 37 | 601 | 0.994 | 0.952 | 0.986 | 0.792 |
| Bits | 31 | 534 | 0.9 | 0.824 | 0.91 | 0.524 |
| Deep Socket | 144 | 1981 | 0.916 | 0.791 | 0.86 | 0.698 |
| Extension Bar | 178 | 500 | 0.951 | 0.914 | 0.955 | 0.735 |
| Flex Handle | 13 | 13 | 0.805 | 0.846 | 0.844 | 0.66 |
| Rachet-Handels | 162 | 200 | 0.943 | 0.908 | 0.945 | 0.764 |
| S Handle | 4 | 4 | 0.771 | 1.0 | 0.995 | 0.835 |
| Sockets | 204 | 3961 | 0.96 | 0.902 | 0.937 | 0.723 |
| Spanners | 1 | 12 | 0.977 | 1.0 | 0.995 | 0.823 |
| long bit socket | 19 | 320 | 0.956 | 0.957 | 0.974 | 0.736 |
| universal joint | 133 | 181 | 0.958 | 0.906 | 0.958 | 0.754 |

**S1.1.3 YOLOV11**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Figure S5.** Confusion Matrix of YOLOV11

**A group of graphs showing the growth of a number of individuals

AI-generated content may be incorrect.**

**Figure S6.** Precision, Recall, Accuracy & loss Graphs of YOLOV11

**Table S3.** Classification Report of YOLOV11

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Class | Images | Instances | Precision (P) | Recall (R) | mAP@0.5 | mAP@0.5:0.95 |
| all | 284 | 8369 | 0.933 | 0.889 | 0.931 | 0.76 |
| Adaptor | 34 | 48 | 0.856 | 0.854 | 0.892 | 0.738 |
| Allen Key | 6 | 6 | 0.903 | 1.0 | 0.995 | 0.905 |
| Bit Holder | 7 | 8 | 1.0 | 0.584 | 0.804 | 0.635 |
| Bit Sockets | 37 | 601 | 0.998 | 0.923 | 0.98 | 0.775 |
| Bits | 31 | 534 | 0.95 | 0.813 | 0.912 | 0.532 |
| Deep Socket | 144 | 1981 | 0.914 | 0.808 | 0.863 | 0.715 |
| Extension Bar | 178 | 500 | 0.95 | 0.934 | 0.954 | 0.759 |
| Flex Handle | 13 | 13 | 0.761 | 0.846 | 0.829 | 0.746 |
| Rachet-Handels | 162 | 200 | 0.933 | 0.899 | 0.937 | 0.78 |
| S Handle | 4 | 4 | 0.916 | 1.0 | 0.995 | 0.947 |
| Sockets | 204 | 3961 | 0.958 | 0.905 | 0.944 | 0.734 |
| Spanners | 1 | 12 | 0.97 | 1.0 | 0.995 | 0.851 |
| long bit socket | 19 | 320 | 0.966 | 0.964 | 0.978 | 0.752 |
| universal joint | 133 | 181 | 0.982 | 0.92 | 0.959 | 0.778 |

**S1.1.4 Faster R-CNN**

**A screenshot of a graph

AI-generated content may be incorrect.**

**Figure S7.** Confusion Matrix of Faster R-CNN

**Table S4.** Classification Report of Faster R-CNN

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Precision | Recall | F1-Score | Support |
| Adaptor | 0.8 | 0.58 | 0.67 | 48 |
| Allen Key | 1.0 | 0.83 | 0.91 | 6 |
| Bit Holder | 0.0 | 0.0 | 0.0 | 8 |
| Bit Sockets | 0.96 | 0.87 | 0.91 | 601 |
| Bits | 0.8 | 0.65 | 0.72 | 534 |
| Deep Socket | 0.9 | 0.83 | 0.87 | 1981 |
| Extension Bar | 0.82 | 0.91 | 0.86 | 500 |
| Flex Handle | 0.77 | 0.77 | 0.77 | 13 |
| Rachet-Handels | 0.85 | 0.89 | 0.87 | 200 |
| S Handle | 0.0 | 0.0 | 0.0 | 4 |
| Sockets | 0.92 | 0.91 | 0.91 | 3961 |
| Spanners | 1.0 | 1.0 | 1.0 | 12 |
| long bit socket | 0.89 | 0.85 | 0.87 | 320 |
| universal joint | 0.84 | 0.91 | 0.87 | 181 |

**S1.2 Image Segmentation**

**S1.2.1 YOLOV5**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Figure S8.** Confusion Matrix of YOLOV5

**A graph of a function

AI-generated content may be incorrect.**

**Figure S9.** Precision, Recall, Accuracy & loss Graphs of YOLOV11

**Table S5.** Classification Report of YOLOV5

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Class | Images | Instances | Box P | Box R | Box mAP50 | Box mAP50-95 | Mask P | Mask R | Mask mAP50 | Mask mAP50-95 |
| all | 287 | 6086 | 0.964 | 0.895 | 0.943 | 0.719 | 0.932 | 0.863 | 0.9 | 0.546 |
| Adaptor | 287 | 46 | 0.987 | 0.739 | 0.86 | 0.615 | 0.903 | 0.674 | 0.748 | 0.39 |
| Allen Key | 287 | 8 | 0.964 | 0.875 | 0.876 | 0.723 | 0.966 | 0.875 | 0.876 | 0.605 |
| Bit Holder | 287 | 17 | 0.97 | 0.941 | 0.947 | 0.771 | 0.973 | 0.941 | 0.947 | 0.704 |
| Bit Socket | 287 | 307 | 0.989 | 0.954 | 0.966 | 0.747 | 0.956 | 0.922 | 0.943 | 0.563 |
| Bits | 287 | 104 | 0.926 | 0.957 | 0.976 | 0.596 | 0.782 | 0.808 | 0.794 | 0.302 |
| Deep Socket | 287 | 2131 | 0.922 | 0.946 | 0.94 | 0.787 | 0.913 | 0.93 | 0.926 | 0.598 |
| Extension Bar | 287 | 281 | 0.98 | 0.891 | 0.972 | 0.742 | 0.953 | 0.864 | 0.934 | 0.559 |
| Flex Handle | 287 | 36 | 1.0 | 0.985 | 0.995 | 0.808 | 1.0 | 0.984 | 0.995 | 0.69 |
| Long bit socket | 287 | 207 | 0.896 | 0.816 | 0.895 | 0.617 | 0.851 | 0.773 | 0.818 | 0.421 |
| Rachet Handle | 287 | 138 | 0.975 | 0.928 | 0.983 | 0.831 | 0.979 | 0.928 | 0.971 | 0.747 |
| S Handle | 287 | 56 | 1.0 | 0.685 | 0.899 | 0.603 | 0.974 | 0.662 | 0.864 | 0.485 |
| Sockets | 287 | 2672 | 0.961 | 0.914 | 0.95 | 0.733 | 0.918 | 0.868 | 0.905 | 0.471 |
| universal joint | 287 | 83 | 0.965 | 1.0 | 0.995 | 0.779 | 0.954 | 0.988 | 0.984 | 0.561 |

**S1.2.2 YOLOV8**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Figure S10.** Confusion Matrix of YOLOV8

**A group of graphs showing the size of a number

AI-generated content may be incorrect.**

**Figure S11.** Precision, Recall, Accuracy & loss Graphs of YOLOV8

**Table S6.** Classification Report of YOLOV8

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Class | Images | Instances | Box P | Box R | Box mAP50 | Box mAP50-95 | Mask P | Mask R | Mask mAP50 | Mask mAP50-95 |
| all | 287 | 6086 | 0.954 | 0.949 | 0.965 | 0.841 | 0.947 | 0.942 | 0.962 | 0.715 |
| Adaptor | 33 | 46 | 0.975 | 0.863 | 0.931 | 0.774 | 0.975 | 0.863 | 0.931 | 0.603 |
| Allen Key | 6 | 8 | 0.871 | 0.846 | 0.87 | 0.792 | 0.871 | 0.845 | 0.87 | 0.582 |
| Bit Holder | 14 | 17 | 0.937 | 1.0 | 0.992 | 0.872 | 0.937 | 1.0 | 0.992 | 0.768 |
| Bit Socket | 24 | 307 | 0.987 | 0.956 | 0.979 | 0.848 | 0.976 | 0.946 | 0.971 | 0.741 |
| Bits | 14 | 104 | 0.938 | 1.0 | 0.994 | 0.753 | 0.911 | 0.971 | 0.989 | 0.684 |
| Deep Socket | 147 | 2131 | 0.955 | 0.955 | 0.957 | 0.857 | 0.952 | 0.953 | 0.952 | 0.769 |
| Extension Bar | 138 | 281 | 0.957 | 0.96 | 0.982 | 0.866 | 0.95 | 0.952 | 0.973 | 0.722 |
| Flex Handle | 33 | 36 | 0.95 | 1.0 | 0.994 | 0.894 | 0.951 | 1.0 | 0.994 | 0.758 |
| Long bit socket | 17 | 207 | 0.932 | 0.923 | 0.961 | 0.816 | 0.913 | 0.903 | 0.95 | 0.683 |
| Rachet Handle | 128 | 138 | 0.982 | 0.971 | 0.988 | 0.935 | 0.982 | 0.971 | 0.991 | 0.823 |
| S Handle | 52 | 56 | 0.985 | 0.929 | 0.939 | 0.83 | 0.966 | 0.911 | 0.937 | 0.681 |
| Sockets | 170 | 2672 | 0.976 | 0.937 | 0.969 | 0.813 | 0.967 | 0.928 | 0.961 | 0.71 |
| universal joint | 81 | 83 | 0.961 | 1.0 | 0.989 | 0.889 | 0.961 | 1.0 | 0.989 | 0.768 |

**S1.2.3 YOLOV11**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Figure S12.** Confusion Matrix of YOLOV11

**A group of graphs showing the size of a number

AI-generated content may be incorrect.**

**Figure S13.** Precision, Recall, Accuracy & loss Graphs of YOLOV11

**Table S7.** Classification Report of YOLOV11

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Class | Images | Instances | Box P | Box R | Box mAP50 | Box mAP50-95 | Mask P | Mask R | Mask mAP50 | Mask mAP50-95 |
| all | 287 | 6086 | 0.978 | 0.953 | 0.973 | 0.859 | 0.971 | 0.951 | 0.969 | 0.734 |
| Adaptor | 33 | 46 | 0.975 | 0.864 | 0.95 | 0.8 | 0.952 | 0.859 | 0.943 | 0.688 |
| Allen Key | 6 | 8 | 0.937 | 0.875 | 0.886 | 0.784 | 0.93 | 0.875 | 0.886 | 0.672 |
| Bit Holder | 14 | 17 | 0.985 | 1.0 | 0.995 | 0.9 | 0.984 | 1.0 | 0.995 | 0.789 |
| Bit Socket | 24 | 307 | 0.997 | 0.96 | 0.98 | 0.86 | 0.99 | 0.955 | 0.974 | 0.754 |
| Bits | 14 | 104 | 0.944 | 0.966 | 0.989 | 0.783 | 0.936 | 0.971 | 0.989 | 0.658 |
| Deep Socket | 147 | 2131 | 0.972 | 0.961 | 0.962 | 0.881 | 0.97 | 0.96 | 0.96 | 0.792 |
| Extension Bar | 138 | 281 | 0.989 | 0.95 | 0.984 | 0.87 | 0.984 | 0.947 | 0.975 | 0.743 |
| Flex Handle | 33 | 36 | 1.0 | 0.994 | 0.995 | 0.927 | 1.0 | 0.997 | 0.995 | 0.766 |
| Long bit socket | 17 | 207 | 0.979 | 0.947 | 0.978 | 0.836 | 0.974 | 0.947 | 0.971 | 0.682 |
| Rachet Handle | 128 | 138 | 0.995 | 0.978 | 0.994 | 0.955 | 0.994 | 0.978 | 0.994 | 0.839 |
| S Handle | 52 | 56 | 0.993 | 0.964 | 0.97 | 0.843 | 0.971 | 0.946 | 0.952 | 0.675 |
| Sockets | 170 | 2672 | 0.984 | 0.939 | 0.973 | 0.833 | 0.976 | 0.933 | 0.967 | 0.73 |
| universal joint | 81 | 83 | 0.965 | 0.996 | 0.991 | 0.9 | 0.965 | 0.997 | 0.991 | 0.761 |

**S1.2.4 Mask -R-CNN**

**A screenshot of a graph

AI-generated content may be incorrect.**

**Figure S14.** Confusion Matrix of Mask R-CNN

**Table S8.** Classification Report of Mask R-CNN

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Precision | Recall | F1-Score | Support |
| Adaptor | 0.45 | 0.27 | 0.33 | 319 |
| Allen Key | 0.0 | 0.0 | 0.0 | 42 |
| Bit Holder | 0.39 | 0.35 | 0.37 | 179 |
| Bit Socket | 0.77 | 0.7 | 0.73 | 2567 |
| Bits | 0.9 | 0.69 | 0.79 | 685 |
| Deep Socket | 0.81 | 0.9 | 0.85 | 20804 |
| Extension Bar | 0.73 | 0.61 | 0.66 | 2711 |
| Flex Handle | 0.31 | 0.33 | 0.32 | 459 |
| Long bit socket | 0.61 | 0.47 | 0.53 | 1399 |
| Rachet Handle | 0.79 | 0.56 | 0.65 | 1463 |
| S Handle | 0.75 | 0.36 | 0.48 | 707 |
| Sockets | 0.89 | 0.89 | 0.89 | 25752 |
| Spanners | 0.47 | 0.06 | 0.11 | 108 |
| universal joint | 0.81 | 0.57 | 0.67 | 971 |

**S2 Object Detection Results**

**A collage of different tools

AI-generated content may be incorrect.**

**Figure S15.** Detection Results

**S3 Image Segmentation Results**

**A close-up of a tool box

AI-generated content may be incorrect.**

**Figure S16.** Segmentation Results

S4. Evaluation Indices

To evaluate the performance of the model various performance metrics such as accuracy, precision, recall (sensitivity or True Positive Rate), and f1 score are calculated based on the confusion matrix.

1. Accuracy: is a measure of total correct prediction done by the model. It is calculated as:

(S1)

1. Precision: it indicates how many positive samples are indeed predicted as positive (how many predicted positives are true). The formula is given as:

(S2)

1. Recall: it is also known as (TPR) or Sensitivity which is the ability of the model to capture all the positive cases (completeness of capturing all positive cases). It is given as:

(S3)

1. F1-Score: is the harmonic mean of the precision and recall. The range is between 1 (the best precision and recall) and 0 (the worst precision and recall). And the formula to calculate the F1 score is given as:

(S4)

**S5. Hyperparameters Optimization**

To ensure reproducibility, all models were trained using the same dataset split and augmentation strategy described in Section 2.1, under identical computational conditions. Table S1 lists the key hyperparameters for the YOLOv11 configuration, which was identified as the optimal model in this study.

Table S9. Training Hyperparameters for YOLOv11

|  |  |
| --- | --- |
| **Parameter** | **Value** |
| Optimizer | SGD with momentum 0.937 |
| Initial Learning Rate | 0.001 |
| Learning Rate Scheduler | Cosine Annealing |
| Batch Size | 16 |
| Training Epochs | 300 |
| Weight Decay | 0.0005 |
| Image Size | 640 × 640 |
| Warm-up Epochs | 3 |
| Loss Function | CIoU loss for bounding box regression, BCE loss for classification |
| Data Augmentation | Mosaic augmentation, random rotation (±10°), horizontal/vertical flip, HSV adjustment |

This configuration was applied consistently to the other YOLO variants for a fair comparison, with only model-specific architectural changes differing between experiments. The same optimizer, learning rate policy, and augmentation techniques were used for Faster R-CNN and Mask R-CNN, with adjustments only to accommodate their specific training requirements.