**8.Sign Language Recognition: A Deep**

There Are Some Important Sub-Areas For Hand Sign Language Recognition Which Are: Hand Detection, Hand Pose Estimation, Real-Time Hand Tracking, Hand Gesture Recognition, And Hand Pose Recovery. Table 4, Table 5, Table 6, Table 7, Show The Details Of These Models. They Are Presented Based On The Same Datasets, Evaluation Metrics, Features, And Input Modalities. The Region-Based Methods Do Not Consider The Whole Image And Work On The Regions Of The Input Image To Localize The Objects. You Only Look Once (Yolo) Uses Just One Cnn To Predict The Bounding Boxes And The Class Probabilities For These Boxes.

To Improve The Detection Accuracy Of Yolo, Another Object Detection Model Has Been Proposed. Provided A Cnn-Based Model For Hand Detection And Segmentation Using Both Of The Unlabeled And Synthetic Data. They Integrated The Features From Multiple Layers Of A Cnn Model To Have A Multi-Scale Representation Of Hand Objects. The Evaluation Results On The Synthetic Data Showed The Improvement Of The Segmentation And Detection Accuracy Obtaining A Detection Accuracy Of 82.0%. Hand Pose Estimation Has Matured Rapidly Due To The Introduction Of Depth Sensors In Recent Years.

Heavy Occlusion, Low Resolution, Varying Illumination Conditions, Different Hand Gestures, And Complex Interactions Between Hands And Objects Are Some Of The Most Important Challenges. There Are A Lot Of Impressive Models For Face And Person Detection That Fusing The Features Of Them By Hand Features May Improve The Detection Accuracy. Designed An End-To-End Approach To Estimate 3d Hand Pose From Stereo Cameras. They Developed A Framework Using 2d Key point Regressor To Estimate The Sparse Disparity Of The Hand Joints. Evaluation Results Show The Model Outperformed State-Of-The-Art Alternatives In Hand Pose Estimation.

A 3d Neural Network Architecture Has Been Provided By Deng Et Al. To 3d Hand Pose Estimation From A Single Depth Image. This Model Achieved The Estimation Error Of 10 Mm And 20 Mm On Icvl And Nyu Datasets. Different Training Strategies Have Been Used To Improve The Performance Of The Joints Localization. Deep Learning Approaches Have Been Proposed To Improve Estimation Accuracy Of Hand Pose. Ge Et Al. Proposed A Multi-View Cnn-Based Model To Project The Query Depth Image Onto Three Orthogonal Planes And Regress The 2d Heat-Maps Of Each Plan In Order To Estimate The Hand Joint Positions.

Fang And Lei Suggested A Cnn Model With An Embedding Auto-Encoder In The Bottom Layer Of The Network To Hand Pose Estimation. Yuan Et Al. Proposed A Tracking System Using The Magnetic Sensors And Inverse Kinematics To Automatically Acquire The Hand Joints Annotations From A Depth Map. Applied A Gan For Hand Pose Estimation By Making A One To One Relation Between Depth Disparity Maps And 3d Hand Pose Models. This Model Refines The Initial Skeleton Estimations For Further Accuracy Improvement.

Results On Msra Show That The Proposed Model Outperforms State-Of-The-Art Methods With A Relative Improvement Of 3 Mm. Double Stream Nets Perform Very Similarly To The Net Trained Only With Depth Images For Hand Pose Estimation. While The Rgb And Depth Images Are Fed Into Two Separate Cnns For Feature Extraction, The Intermediate Layers Of The Cnns Are Fused For Final Estimation. The Impressive Capability Of Cnn To Work With The Still Images Is Not Enough For Video Inputs. Hands Are The Most Important Object In The Inputs Of The Sign Language Recognition Models.

Tracking The Detected Hands Is One Of The Substantial Challenges For Video Inputs Due To High Occlusions Of Hand Fingers And Joints. We Review The Deep-Based Suggested Models For Hand Tracking In The Last Four Years. Proposed A Real-Time Finger-Spelling Recognition Model Using Cnn From The Depth Map. They Have Recorded Some Depth Videos In Multiple Subjects And Used Different Learning Configurations To Classify The Alphabets And Numbers. The Evaluation Results On Their Dataset Have Led To The Accuracy Of 0.99 Percent.

Hand Gesture Recognition Area Can Provide Fundamental Information For Hci Applications. Some Models Consider Only A Uniform Or Static Background, Avoid The Rapid Hand Motions, Or Assume The Hand As A Skin-Colored Object Presented A Dynamic Graph-Based Spatial–Temporal Attention (Dg-Sta) Model For Hand Gesture Recognition. Results Of This Model On Skig, Natops, And Sbu Datasets Demonstrate A Relative State-Of-The-Art Recognition Accuracy Improvement Of 0.19 Percent, 8.52 Percent, And 4.11 Percent.

Canuto Dos Santos Developed A Deep-Based Model Using Two Resnet Models And A Soft-Attention Ensemble Layer For Dynamic Gestures Provided A Convolutional Two-Stream Consensus Voting Network To Explicitly Model The Short-Term And Long-Term Structure Of The Rgb Sequences. Their Model Out Performed The State-Of-The-Art Models On These Datasets With A Relative Improvement Of 4.47 Percent And 0.61 Percent. Have Proposed A Hand Sign Recognition Model Using Rbm From Visual Data. Their Model Has The Capability Of Multiple Gestures Generalization For Only One Person Or Multiple Person. Results Showed That The Model Achieved State-Of-The-Art With A Relative Accuracy Improvement Of 27.31 Percent.