When you are creating training sets for deep learning, there are a lot of different ways to go about labeling images and the contents of them for the training algorithm to learn from. The explosive development of deep learning started from an ImageNet competition where participants were tasked with creating an algorithm to correctly label images. From there, many developments were made in extremely short periods of time compared to how it started. Dense predictions, which are predictions made for every individual pixel with an image or video segment, were created. This created the entry for predicting different layers of an image. Currently, supervised techniques, which are trained by people manually labeling images are performing the best in many metrics. Weakly-supervised techniques, which train from incomplete annotations, are significantly behind supervised techniques. Semi-automatic techniques, where humans are required for training are another way to avoid expensive labeling, however they require interaction at testing time, which comes in the forms of: scribbles or bounding boxes, which basically surround the object of interest. Deep Extreme Cuts, also known as DEXTR, obtains the extreme points: the left, right, top, and bottom-most pixels of the object to then allow for a more accurate annotation. DEXTR can also incorporate additional boundary points past the extreme ones to allow better identification. However, some methods find that their performance varies greatly from training sets to test sets. This isn't the case when it comes to DEXTR, the aglorithsm that trained on DEXTR annotation algorithms, performed as good as when trained from ground truth annotations. For a given target quality, for example a set accuracy, training with DEXTR is more efficient than training with ground truth annotations. Different methods of segmentation from points include: click carving, where you can interactively update the results of video object segmentation using user-defined clicks, iFCN which guides annotation using positive and negative points, and grabcut which creates a box around the object of interest. DEXTR greatly improves upon the results by adding the 4 extreme points as mentioned before as supervisory signals. Popular grouping methods provide instance segmentation in the form of automatically segmented object proposals, while other variants provide instance level segmentation from weak guiding signals. Both group methods accuracy has increassed vastly recently, when using recent strong supervisory signals such as extreme points or bounding boxes. Now to talk more about extreme points. A common weakly-supervissed approach is drawing a bounding box. This method requires a user to click points outside of the object and drag a box tightly around the object. DEXTR avoids possible errors by finding the outermost points in each direction, which allows for the tightest fit possible that always includes all of the edges of an object. Using extreme points to map objects, we can also employ a heatmap with activations in the regions of the box created using extreme points. The performance using DEXTR on testing data and training data was very similar, with minimal difference whether or not the class had been seen during the training or not. DEXTR trains on PASCAL 2012 segmentation for 100 epochs, or on COCO 2014 training set for 10 epochs. The training with listed parameters, took 20 hours using PASCAL and 5 days using COCO. Using Ablation studies, where parts of the neural network are removed to see the impact of every individual part, we can observe how different parts of the neural network are important to reaching the final product. The team tested each individual component of the DEXTR method across many popular databases, and showed great results on all of them.