We propose EXTENT, which combines the context and content of an image to provide the image with metadata that cannot be inferred by either context or content alone. EXTENT first applies contextual analysis for reducing the complexity of content analysis, it can then afford using more expensive (hence robust) algorithms for performing content analysis. Our experiments show that effectively combining content with context can infer metadata accurately.