This paper presents an automated traffic safety diagnostics solution using deep learning techniques to process traffic videos by Unmanned Aerial Vehicle (UAV). Mask R-CNN is employed to better detect vehicles in UAV videos after video stabilization. The vehicle trajectories are generated when tracking the detected vehicle by Channel and Spatial Reliability Tracking (CSRT) algorithm. During the detection process, missing vehicles could be tracked by the process of identifying stopped vehicles and comparing Intersect of Union (IOU) between the tracking results and the detection results. In addition, rotated bounding rectangles based on the pixel-"to-pixel" manner masks that are generated by Mask R-CNN detection, which are also introduced to obtain precise vehicle size and location data. Moreover, surrogate safety measures (i.e. post-encroachment time (PET)) are calculated for each conflict event at the pixel level. Therefore, conflicts could be identified through the process of comparing the PET values and the threshold. To be more specific, conflict types that include rear-end, head-on, sideswipe, and angle could be determined. A case study is presented at a typical signalized intersection, the results indicate that the proposed framework could notably improve the accuracy of the output data. Furthermore, by calculating the PET values for each conflict event, an automated traffic safety diagnostic for the studied intersection could be conducted. According to the research, rear-end conflicts are the most prevalent conflict type at the studied location, while one angle collision conflict is identified at the study duration. It is expected that the proposed method could help diagnose the safety problems efficiently with UAVs and appropriate countermeasures could be proposed after then.