

We present Vision-based Navigation with Language-based Assistance (VNLA), a grounded vision-language task where an agent with visual perception is guided