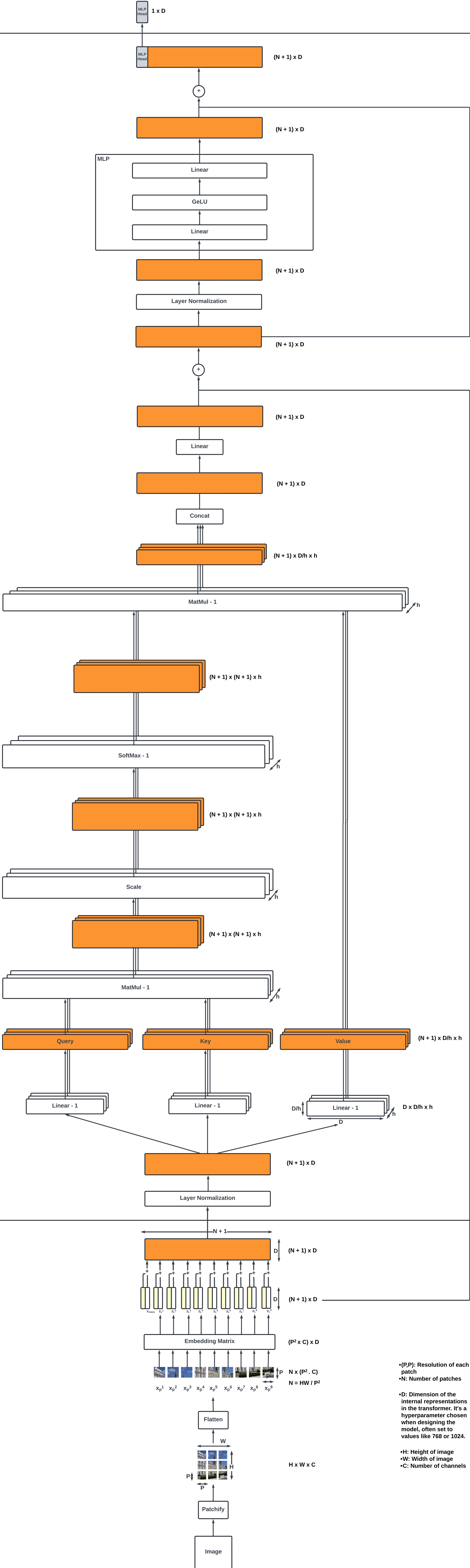


L x



•(P,P): Resolution of each patch  
•N: Number of patches

•D: Dimension of the internal representations in the transformer. It's a hyperparameter chosen when designing the model, often set to values like 768 or 1024.

•H: Height of image  
•W: Width of image  
•C: Number of channels