

https://kazemnejad.com/blog/transformer_architecture_positional_encoding/
<http://jalammar.github.io/illustrated-transformer/>
<https://leemeng.tw/neural-machine-translation-with-transformer-and-tensorflow2.html>
<https://www.youtube.com/watch?v=ugWDII0HtPA&t=2295s>
<https://www.youtube.com/watch?v=S27pHKBEp30&t=772s>
<https://www.youtube.com/watch?v=spUNpyF58BY&t=2s>
[http://speech.ee.ntu.edu.tw/~tlkagk/courses/ML_2019/Lecture/Transformer%20\(v5\).pdf](http://speech.ee.ntu.edu.tw/~tlkagk/courses/ML_2019/Lecture/Transformer%20(v5).pdf)
<https://blog.christianperone.com/2017/11/the-effective-receptive-field-on-cnns/>
<https://towardsdatascience.com/transformers-141e32e69591>
<https://numpy.org/devdocs/user/theory.broadcasting.html>
https://www.youtube.com/watch?v=l_3zj6HeWUE