Title	Results
The Expressive Power of Neural Networks: A View from the Width (2017)	Any Lebesgue-integrable function f from R^n to R can be approximated by a fully-connected width-(n + 4) ReLU network to arbitrary accuracy with respect to L_1 distance. For integer k, there exist a class of width-O(k^2) and depth-2 ReLU networks that cannot be approximated by any width-O($k^{1.5}$) and depth-k networks.
The Power of Depth for Feedforward Neural Networks (2016)	The existence of a 3-layer network, which cannot be realized by any 2-layer to more than a constant accuracy if the size is subexponential in the dimension.
Deep Network Approximation for Smooth Functions (2020)	ReLU forward neural networks (FNN) with width O(Nln$N$) and depth O(Lln$L$) can approximate functions in the unit ball of C^s ($[0,1]^d$) with approximation rate O($N^{-2s/d}L^{-2s/d}$).
Why deep neural networks for function approximation (2017)	In order to approximate a function which is $\Theta(\log(1/\varepsilon))$-order derivable with ε error universally, a deep network with O($\log(1/\varepsilon)$) layers and O(poly $\log(1/\varepsilon)$) weights can do but Ω(poly $1/\varepsilon$) weights will be required if there is only o(log $1/\varepsilon$) layers.
Error bounds for approximations with deep ReLU networks (2017)	C^n -functions on R^d with a bounded domain can be approximated with ε error universally by a ReLU network with O($\log(\frac{1}{\varepsilon})$) layers and O($(\frac{1}{\varepsilon})^{d/n}\log(\frac{1}{\varepsilon})$)weights.
Representation Benefits of Deep Feedforward Networks (2015)	For any integer k explicitly constructed networks with O(k^3) layers and constant width which cannot be realized by any network with O(k) layers whose size is smaller than 2^k.
On the Expressive Power of Deep Learning: A Tensor Analysis (2016)	proved the existence of classes of deep convolutional ReLU networks that cannot be realized by shallow ones if its size is no more than an exponential bound.
Optimal Approximation of Continuous Functions by Very Deep ReLU Networks (2018)	We prove that constant-width fully-connected networks of depth provide the fastest possible approximation rate that cannot be achieved with less deep networks
Exponential Convergence of the Deep Neural Network Approximation for Analytic Functions (2018)	Deep ReLU networks can approximate functions of d variables as well as linear approximation by algebraic polynomials with a comparable number of parameters.
Deep Network Approximation Characterized by Number of Neurons (2020)	It is shown by construction that ReLU FNNs with width C_1 $\max\{d\lfloor N^{\frac{1}{d}}, N+1\rfloor\}$ and depth $12L + C_2$can approximate an arbitrary Holder continuous function of order α with a constant λ on $[0,1]^d$ with a nearly tight approximation rate $19\sqrt{d}\lambda N^{-2\alpha/d}L^{-2\alpha/d}$measured in L_p-norm for any given $N, L \in N^+$ and $p \in [1,\infty]$.

Nonlinear Approximation and (Deep) ReLU Networks (2021)	A theoretical result that the approximation of univariate functions by ReLU networks possesses greater approximation power than these traditional methods of nonlinear approximation such as variable knot splines or n-term approximation from dictionaries. This result corroborates with empirical results that large classes of functions can be efficiently captured by neural networks where classical nonlinear methods fall short of the task.