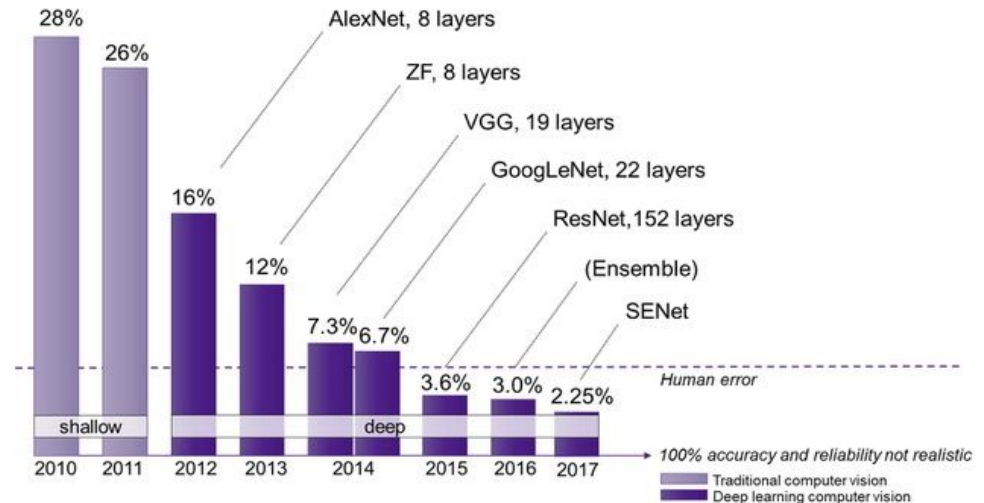
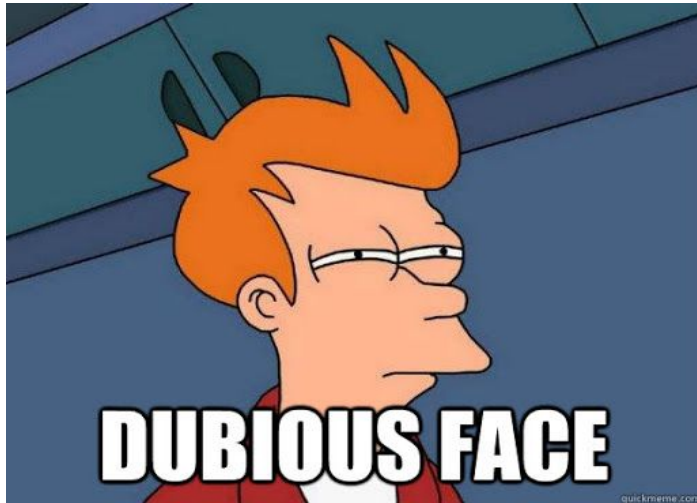


- Machine's ImageNet classification accuracy is higher than humans
- Yet none of us believe that machines are superior than humans
- We suspect this mismatch comes from the evaluation protocol



Adversarial Examiner (AE)

- **Dynamic test set based on test history** instead of fixed test set
- **Worst case** instead of Average case

