

Ước lượng kênh dựa trên phương pháp học sâu

Trương Đức Trường - 20210880

Trường Điện - Điện tử,

Đại học Bách Khoa Hà Nội,

Số 1 Đại Cồ Việt, Hai Bà Trưng, Hà Nội.

Email: truong.td210880@sis.hust.edu.vn

Nguyễn Minh Hiếu - 20210352

Trường Điện - Điện tử,

Đại học Bách Khoa Hà Nội,

Số 1 Đại Cồ Việt, Hai Bà Trưng, Hà Nội.

Email: hieu.nm210352@sis.hust.edu.vn

Abstract—Trong bài báo cáo chúng em tìm hiểu một thuật toán dựa trên học sâu để ước lượng kênh trong các hệ thống thông tin. Đầu tiên từ việc coi đáp ứng thời gian–tần số của một kênh thông tin liên lạc như một hình ảnh 2D. Mục tiêu là tìm các giá trị chưa biết của đáp ứng kênh bằng cách sử dụng một số giá trị đã biết tại các vị trí pilot. Để làm được điều này, ta sử dụng các kỹ thuật deep image là image super-resolution (SR) và image restoration (IR). Quy trình này xem các giá trị pilot như một hình ảnh có độ phân giải thấp và sử dụng một mạng SR kết hợp với một mạng IR khử nhiễu để ước lượng kênh. Loss dựa trên MSE cho thấy thuật toán có thể so sánh với phương pháp MMSE (Minimum Mean Square Error) khi có đầy đủ thông tin thống kê kênh. Kết quả đo được cho thấy rằng phương pháp này có thể sử dụng hiệu quả trong ước lượng kênh truyền.

I. GIỚI THIỆU

Phương pháp điều chế ghép kênh phân chia tần số trực giao (OFDM) đã được sử dụng rộng rãi trong các hệ thống thông tin liên lạc để giải quyết hiện tượng fading chọn lọc tần số trong các kênh không dây. Trong một kênh thông tin liên lạc, tín hiệu nhận được thường bị méo do đặc tính của kênh. Để khôi phục các ký hiệu được truyền, ảnh hưởng của kênh phải được ước lượng và bù trừ tại bộ thu.

Thông thường, bộ thu ước lượng kênh bằng cách sử dụng một số ký hiệu được gọi là "pilot", vị trí và giá trị của chúng trong miền thời gian và miền tần số được cả bên phát và bên thu biết trước. Tùy thuộc vào cách sắp xếp các ký hiệu pilot [1], có ba cấu trúc khác nhau: loại khối (block-type), loại lược (comb-type), và loại lưới (lattice-type).

- Trong sắp xếp loại khối, các pilot được truyền định kỳ ở đầu mỗi khối OFDM tại tất cả các sóng mang con.
- Trong sắp xếp loại lược, các pilot chỉ xuất hiện ở một số sóng mang con trong một số ký hiệu OFDM.
- Trong sắp xếp loại lưới, các pilot được chèn dọc theo cả trục thời gian và tần số theo một chòm sao hình thoi.

Các phương pháp ước lượng kênh dựa trên pilot truyền thống như Least Square (LS) và Minimum Mean Square Error (MMSE) [2] sử dụng các giá trị pilot trong miền thời gian và miền tần số để tìm đáp ứng kênh truyền. Các thuật toán này đã được tối ưu hóa trong nhiều điều kiện khác nhau. Ngược lại với phương pháp LS, không yêu cầu thông tin về thống kê của kênh, phương pháp MMSE đạt hiệu suất tốt hơn bằng cách sử dụng thống kê của kênh và phương sai nhiễu.

Gần đây, kỹ thuật học sâu (Deep Learning - DL) được sử dụng nhiều trong các hệ thống thông tin liên lạc. Dựa

trên DL, một số phương pháp đã được đề xuất để cải thiện hiệu suất của các thuật toán truyền thống như modulation recognition, signal detection, channel equalization, channel state information (CSI), và channel estimation.

Trong bài viết này, ta cũng sử dụng Deep Learning để ước lượng kênh trong hệ thống OFDM. Trong phương pháp này, miền thời gian và miền tần số của đáp ứng kênh được mô hình hóa như một hình ảnh 2D, chỉ biết tại các vị trí pilot. Đáp ứng kênh truyền tại các vị trí pilot được xem như một hình ảnh độ phân giải thấp và đầu ra kênh ước lượng được xem là một hình ảnh độ phân giải cao. Thuật toán được tiếp cận qua hai giai đoạn để ước lượng kênh truyền:

- Bước đầu tiên sử dụng mạng image super-resolution (SR) với đáp ứng kênh truyền giống như là một hình ảnh đầu vào chất lượng thấp.
- Sau đó sử dụng mạng image restoration (IR) để loại bỏ nhiễu.

Hai thuật toán CNN là SRCNN và DnCNN, đã được sử dụng cho các mạng SR và IR tương ứng.

II. KIẾN THỨC NỀN TẢNG

A. Ước lượng kênh

Trong hệ thống OFDM [3], mối quan hệ giữa tín hiệu đầu vào và đầu ra tại khe thời gian thứ k và sóng mang con thứ i được biểu diễn như công thức dưới đây:

$$Y_{i,k} = H_{i,k}X_{i,k} + Z_{i,k}. \quad (1)$$

Với một khung con OFDM có kích thước $N_S \times N_D$, chỉ số khe thời gian k nằm trong khoảng $[0, N_D - 1]$, và phạm vi của chỉ số sóng mang con i là $[0, N_S - 1]$. Trong phương trình (1), $Y_{i,k}$, $X_{i,k}$ và $Z_{i,k}$ lần lượt là tín hiệu nhận, ký hiệu OFDM được truyền và nhiễu Gaussian trắng tương ứng. $H_{i,k}$ là phần tử (i, k) của ma trận $H \in \mathbb{C}^{N_S \times N_D}$, biểu diễn đáp ứng thời gian–tần số của kênh cho tất cả các sóng mang con và khe thời gian.

Để ước lượng kênh, đặc biệt trong các kênh có fading, đáp ứng miền thời gian được biểu diễn dưới dạng $H = \{h_1, h_2, \dots, h_{N_D}\}$, trong đó mỗi h_k là đáp ứng kênh trong miền tần số tại khe thời gian thứ k .

Phương pháp Least Square (LS) ước lượng kênh tại các vị trí pilot. Giả sử ma trận kênh LS ước lượng tại các vị

trí pilot được biểu diễn dưới dạng một ma trận đường chéo $H_p^{LS} \in \mathbb{C}^{N_P \times N_P}$, thì H_p^{LS} có thể được ước lượng bằng:

$$\hat{H}_p^{LS} = \arg \min_{H_p} \|y_p - H_p x_p\|_2^2 \quad (2)$$

trong đó $\|\cdot\|_2$ là chuẩn Euclid, $\hat{H}_p^{LS} \in \mathbb{C}^{N_P \times N_P}$ là ma trận đường chéo với các giá trị trên đường chéo là ước lượng kênh tại các vị trí pilot. x_p chứa các giá trị pilot đã biết, và y_p là các tín hiệu quan sát tương ứng. Giải phương trình trên ta có kết quả $\hat{h}_{LS} = \text{diag}(\hat{H}_{LS}) = y_p/x_p$. Để tìm các giá trị tại các điểm không phải vị trí của pilot ta sử dụng phương pháp nội suy hai chiều.

Một lựa chọn tốt hơn LS là phương pháp MMSE, được tính bằng cách nhân các giá trị LS tại vị trí các ký hiệu pilot với ma trận $A_{MMSE} \in \mathbb{C}^{N_L \times N_P}$:

$$\hat{h}_d^{MMSE} = A_{MMSE} \times \hat{h}_p^{LS} \quad (3)$$

trong đó $\hat{h}_d^{MMSE} \in \mathbb{C}^{N_L \times 1}$ ($N_L = N_S \times N_D$) là vector ước lượng MMSE của đáp ứng kênh H trong khung con d . Để xác định ma trận lọc, sai số trung bình bình phương (MSE) được tính như sau:

$$\varepsilon = \mathbb{E}\{\|\hat{h}_d - A_{MMSE} \hat{h}_p^{LS}\|_2^2\} \quad (4)$$

Việc tối thiểu hóa ε dẫn đến:

$$A_{MMSE} = R_{h_d h_p} (R_{h_p h_p} + \sigma_n^2 (x x^H)^{-1})^{-1} \quad (5)$$

trong đó ma trận $R_{h_d h_p} = \mathbb{E}\{h_d h_p^H\}$ biểu diễn ma trận tương quan kênh giữa khung con mong muốn và các ký hiệu pilot, và $R_{h_p h_p} = \mathbb{E}\{h_p h_p^H\}$ là ma trận tương quan tại các vị trí pilot. Có thể thấy, phương pháp MMSE chỉ hữu ích khi ma trận tương quan kênh R được biết đầy đủ.

B. Super-Resolution (SR) và Image Restoration (IR)

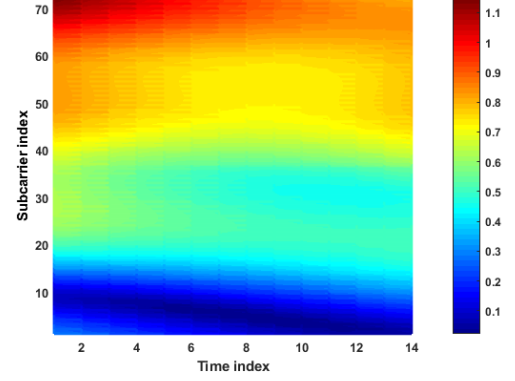
Khi xử lý một hình ảnh có độ phân giải thấp (LR) và chứa nhiễu, nhiều kỹ thuật đã được đề xuất để tái tạo một hình ảnh có độ phân giải cao hơn (HR) và ít nhiễu hơn. Siêu phân giải hình ảnh (SR) là một nhóm các kỹ thuật được sử dụng để nâng cao độ phân giải của hình ảnh. Các thuật toán dựa trên học sâu, đặc biệt là mạng nơ-ron tích chập (CNN), đã đạt được hiệu suất cao trong việc khôi phục hình ảnh HR từ dữ liệu đầu vào LR.

Gần đây, mạng SRCNN [5] (Super-resolution convolutional neural network) được đề xuất để ánh xạ giữa hình ảnh LR và HR theo cách end-to-end. Ngoài ra, các thuật toán phục hồi hình ảnh (IR) cũng có thể được áp dụng để loại bỏ/giảm nhiễu trong hình ảnh. Một ví dụ là mạng DnCNN [6] (Denoising CNN), sử dụng học phần dư và chuẩn hóa để tăng tốc quá trình huấn luyện.

III. MÔ HÌNH HỆ THỐNG

A. Ánh của kênh

Trong báo cáo này, ta sử dụng một liên kết giữa một cặp ăng-ten truyền (Tx) và ăng-ten nhận (Rx), tức là xét liên kết truyền thông single-input, single-output (SISO). Đối với liên kết này, ma trận đáp ứng thời gian-tần số của kênh H (kích



Hình 1. Hình ảnh 2D được chuẩn hóa của một kênh truyền.

thước $N_S \times N_D$) giữa một máy phát và một máy thu. Ma trận H này có các giá trị phức có thể được biểu diễn thành hai hình ảnh 2D (một hình ảnh 2D cho phần thực và một hình ảnh 2D cho phần ảo).

Ví dụ về hình ảnh 2D được chuẩn hóa cho phần thực của một ma trận theo miền thời gian và tần số kênh mẫu với $N_D = 14$ khe thời gian và $N_S = 72$ sóng mang con được hiển thị trong Hình 1.

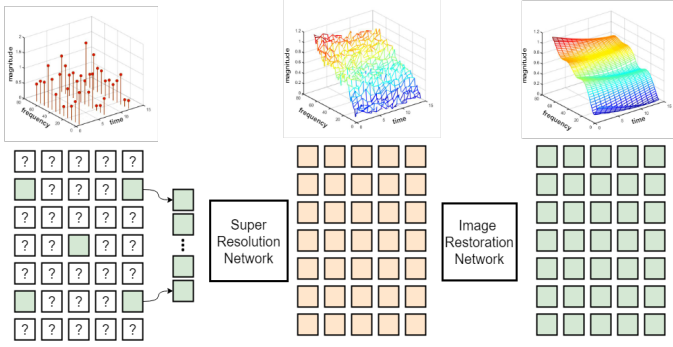
B. Cấu trúc mạng

Tổng quan về quy trình sử dụng để ước lượng kênh được minh họa trong Hình 2. Mục tiêu là ước lượng toàn bộ miền thời gian-tần số của kênh sử dụng các pilot được truyền. Tương tự như tiêu chuẩn LTE, cách sắp xếp pilot dạng lưới (lattice-type) đã được sử dụng để truyền pilot.

Giá trị ước lượng của kênh tại các vị trí pilot \hat{h}_p^{LS} (có thể nhiễu) được coi là phiên bản có độ phân giải thấp (LR) và chứa nhiễu của hình ảnh kênh. Để thu được hình ảnh kênh hoàn chỉnh, quy trình huấn luyện hai giai đoạn:

- Trong giai đoạn đầu, một mạng SR được triển khai, lấy \hat{h}_p^{LS} dưới dạng hình ảnh đầu vào có độ phân giải thấp (được vector hóa, lần lượt cho phần thực và phần ảo) và ước lượng các giá trị chưa biết của đáp ứng kênh H .
- Trong giai đoạn hai, một mạng IR khử nhiễu được ghép nối với mạng SR để loại bỏ các hiệu ứng nhiễu.

Trong các mạng SR và IR, chúng em lần lượt sử dụng SRCNN và DnCNN. SRCNN trước tiên sử dụng một sơ đồ nội suy để tìm các giá trị xấp xỉ của hình ảnh độ phân giải cao (kênh), sau đó cải thiện độ phân giải bằng cách sử dụng một mạng tích chập ba lớp. Lớp tích chập đầu tiên sử dụng 64 bộ lọc kích thước 9×9 , lớp thứ hai sử dụng 32 bộ lọc kích thước 1×1 và cả hai đều kết hợp hàm kích hoạt ReLU, lớp cuối cùng sử dụng một bộ lọc kích thước 5×5 để tái tạo hình ảnh. DnCNN là một mạng học phần dư với 20 lớp tích chập. Lớp đầu tiên sử dụng 64 bộ lọc kích thước $3 \times 3 \times 1$ với hàm kích hoạt ReLU, 18 lớp tiếp theo sử dụng 64 bộ lọc kích thước $3 \times 3 \times 64$, kết hợp batch normalization và ReLU, lớp cuối cùng sử dụng một bộ lọc kích thước $3 \times 3 \times 64$ để tái tạo đầu ra. Hình 2 thể hiện quá trình trên.



Hình 2. Hình ảnh mô phỏng quá trình huấn luyện.

C. Training

Gọi tập hợp tất cả các tham số mạng là $\Theta = \{\Theta_S, \Theta_R\}$, trong đó Θ_S và Θ_R lần lượt là các tham số của mạng SR và IR.

Đầu vào là vector giá trị pilot \hat{h}_p^{LS} , và đầu ra là ma trận kênh ước lượng \hat{H} được định nghĩa như sau:

$$\hat{H} = f(\Theta; \hat{h}_p^{LS}) = f_R(f_S(\Theta_S; \hat{h}_p^{LS}); \Theta_R) \quad (6)$$

với f_S và f_R lần lượt là các hàm SR và IR.

Hàm loss của mạng là sai số trung bình bình phương (MSE) giữa kênh ước lượng và kênh thực, được tính như sau:

$$L = \frac{1}{|T|} \sum_{h_p \in T} \|f(\Theta; \hat{h}_p^{LS}) - H\|_2^2 \quad (7)$$

trong đó T là tập dữ liệu huấn luyện và H là kênh lý tưởng.

Để đơn giản hóa quá trình huấn luyện, ta có thể áp dụng thuật toán huấn luyện theo hai giai đoạn:

- Giai đoạn đầu, tối ưu hóa mạng SR bằng cách tối thiểu hóa hàm mất mát L_1 :

$$L_1 = \frac{1}{|T|} \sum_{h_p \in T} \|Z - H\|_2^2 \quad (8)$$

trong đó $Z = f_S(\Theta_S; \hat{h}_p^{LS})$ là đầu ra của mạng SR.

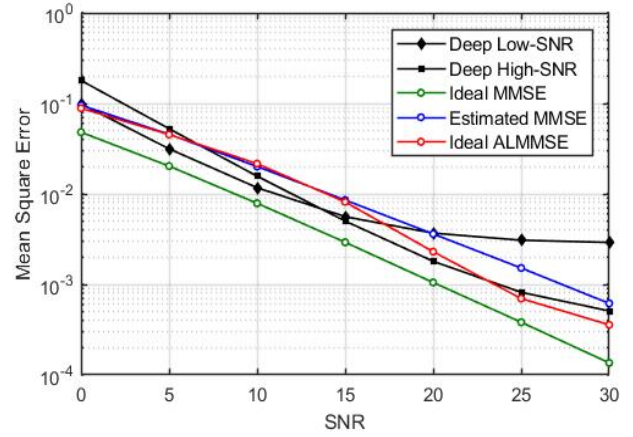
- Giai đoạn hai, giữ cố định trọng số của mạng SR và tối ưu hóa mạng IR bằng cách xác định $\hat{H} = f_R(Z; \Theta_R)$ và tối thiểu hóa hàm mất mát L_2 :

$$L_2 = \frac{1}{|T|} \sum_{h_p \in T} \|\hat{H} - H\|_2^2 \quad (9)$$

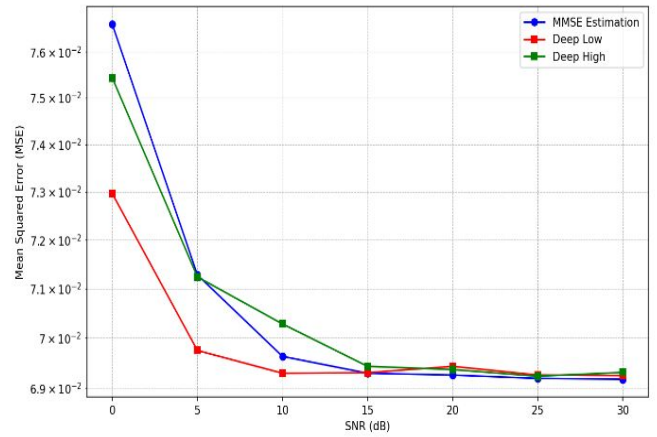
Lưu ý rằng, tại mỗi điều kiện đường truyền khác nhau phụ thuộc vào SNR trọng số của mô hình khác nhau. Vì vậy, để đạt được kết quả toàn diện, mô hình cần được huấn luyện lại cho mỗi giá trị SNR.

IV. KẾT QUẢ MÔ PHỎNG

Dưới đây chúng em mô phỏng lại thuật toán Deep learning với một bên phát và một bên thu trên nhiều dải giá trị SNR và đánh giá sai số bằng phương pháp bình phương tối thiểu (MSE). Em đã sử dụng ngôn ngữ Python để mô phỏng lại quá trình ước lượng kênh của hệ thống OFDM với kênh truyền là



Hình 3. Kết quả MSE cho mô hình kênh VehA.



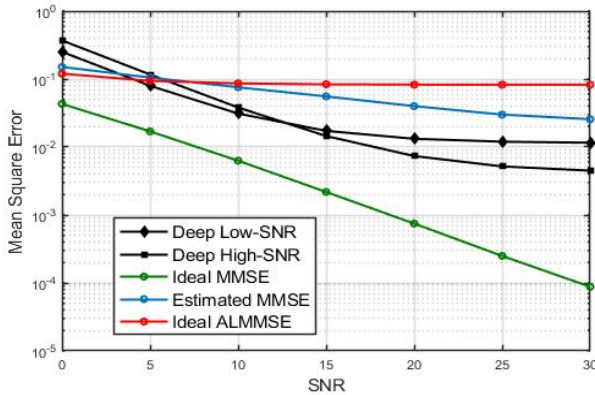
Hình 4. Kết quả MSE đối với mô hình kênh VehA tự triển khai.

VehA. Đối với mô hình Deep Learning, em sử dụng Pytorch để triển khai mô hình và thực hiện huấn luyện mô hình trên máy tính cá nhân sử dụng GPU NVIDIA RTX 3050Ti. Các tập dữ liệu huấn luyện em lấy của tác giả bài [4]. Hình 4 là kết quả mà chúng em đã mô phỏng lại được.

Còn trong [4], người ta thực hiện huấn luyện mô hình và đánh giá sai số trung bình bình phương (MSE) trên một dải giá trị SNR, đồng thời so sánh kết quả với các thuật toán cơ sở phổ biến. Ở đây, họ xem xét một hệ thống truyền thông có một ăng-ten ở phía phát và một ăng-ten ở phía nhận. Mô phỏng sử dụng trình giả lập LTE để mô hình hóa kênh và truyền pilot. Việc triển khai mô hình đề xuất được thực hiện bằng Keras và Tensorflow với GPU làm backend. Đối với mạng SR và IR, learning rate được đặt là 0.001, kích thước batch size là 128 và số lượng vòng lặp tối đa là 500. Các tập training, testing và validation lần lượt gồm 32.000, 4.000 và 4.000 mẫu kênh. Trong các mô phỏng, mỗi khung gồm 14 khe thời gian với 72 sóng mang con (theo chuẩn LTE) trong hai trường hợp là Vehicular-A (VehA) là kênh với độ trễ ngắn và SUI5 là mô hình kênh với độ trễ dài, băng thông 1.6 MHz, tần số sóng mang là 2.1 GHz, và tốc độ thiết bị người dùng (UE)

là 50 km/h. Để đánh giá hiệu suất, ta so sánh độ chính xác của phương pháp với ba thuật toán là : MMSE lý tưởng (idea MMSE); MMSE ước lượng (estimate MMSE); ALMMSE lý tưởng (idea ALMMSE).

Từ Hình 3 ta thấy được MMSE lý tưởng có hiệu suất tốt nhất và cung cấp ngưỡng dưới của MSE. Tuy nhiên, phương pháp này yêu cầu biết đầy đủ ma trận tương quan kênh, điều này không thực tế. Với các giá trị SNR thấp, mô hình được huấn luyện tại SNR = 12 dB (gọi là deep low-SNR) có hiệu suất tương đương với MMSE lý tưởng và vượt trội so với ALMMSE lý tưởng và MMSE ước lượng. Với SNR cao hơn một ngưỡng nhất định, mô hình được huấn luyện tại SNR = 22 dB (deep high-SNR) bắt đầu hoạt động tốt hơn mạng deep low-SNR. Do đó, ta có thể chia phạm vi SNR thành hai vùng. Với SNR thấp, deep low-SNR được sử dụng; khi vượt quá ngưỡng SNR, deep high-SNR được sử dụng. Tuy nhiên, khi SNR lớn hơn 23 dB, hiệu suất của mạng deep high-SNR giảm dần, điều này cho thấy cần huấn luyện thêm một mạng mới cho các giá trị SNR cao hơn. Trong phạm vi SNR dưới 20 dB, hai mạng này là đủ.



Hình 5. Kết quả MSE cho mô hình kênh SUI5.

Từ Hình 5 ta thấy tại SNR trên 5 dB, các phương pháp như ALMMSE và MMSE ước lượng bị suy giảm đáng kể, trong khi mô hình Deep Learning đề xuất vẫn có khả năng khám phá các thống kê tiềm ẩn và đạt được MSE chấp nhận được. MMSE lý tưởng vẫn là phương pháp tốt nhất, nhưng không khả thi trong các ứng dụng thực tế vì cần đầy đủ thông tin thống kê kênh.

V. KẾT LUẬN

Báo cáo này đã đưa ra được mô hình ước lượng kênh sử dụng phương pháp học sâu. Qua bài tập lớn lần này, chúng em đã tiếp thu được thêm nhiều kiến thức về kỹ thuật ước lượng kênh truyền cho hệ thống truyền thông, ứng dụng các kiến thức về Deep Learning, trau dồi kỹ năng đọc tài liệu và viết báo cáo. Mặc dù đã cố gắng, tuy nhiên do giới hạn về mặt kiến thức và kinh nghiệm nên báo cáo của chúng em đã có sự sai sót nhiều so với bài báo. Do đó rất mong đợi được sự quan tâm góp ý của thầy.

Chúng em xin chân thành cảm ơn !

REFERENCES

- [1] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," IEEE Trans.
- [2] Pukkila, M. (2000). Channel estimation modeling. Nokia Research Center, 17, 66.
- [3] Y. Liu, Z. Tan, H. Hu, L. J. Cimini and G. Y. Li, "Channel Estimation for OFDM," in IEEE Communications Surveys and Tutorials, vol. 16, no. 4, pp. 1891-1908, Fourthquarter 2014.
- [4] M.Soltani, V.Pourahmadi, A.Mirzaei and H.Sheikhzadeh, "Deep Learning-Based Channel Estimation," in IEEE Communications Letters, vol.23, no.4, pp.652-655, April 2019. Broadcast., vol. 48, no. 3, pp. 223-229, Sep. 2002
- [5] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 2, pp. 295-307, Feb. 2016.
- [6] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," IEEE Trans. Image Process., vol. 26, no. 7, pp. 3142-3155, Jul. 2017.