

MÔ HÌNH KHỬ NHIỄU ẢNH DỰA TRÊN TỔNG BIẾN PHÂN THÍCH NGHIỆ

AN IMAGE DENOISING MODEL USING AN ADAPTIVE TOTAL VARIATION REGULARIZATION

Phan Trần Đăng Khoa

Trường Đại học Bách khoa – Đại học Đà Nẵng; ptdkhoa@dut.udn.vn

Tóm tắt - Trong bài báo này, chúng tôi trình bày về mô hình khử nhiễu ảnh dựa trên hướng tiếp cận biến phân. Để có kết quả khử nhiễu tốt hơn và bảo toàn đường biên, chúng tôi đề xuất thành phần tổng biến phân thích nghi được xây dựng dựa trên hàm trọng số có giá trị biến đổi tùy thuộc vào đặc trưng của điểm ảnh. Phép đo độ cong trung bình được sử dụng để mô tả đặc trưng của ảnh và điều chỉnh giá trị của hàm trọng số, qua đó thay đổi độ làm mịn ảnh của mô hình. Phương pháp Split Bregman được áp dụng để giải bài toán tối ưu. Kết quả mô phỏng cho thấy, mô hình đề xuất có khả năng khử nhiễu tốt hơn so với mô hình cổ điển đối với cả hai tiêu chí định lượng và định tính.

Từ khóa - khử nhiễu; tổng biến phân; Split Bregman; mô hình ROF

Abstract - In this paper, we present a denoising model based on the variational approach. To better denoise and preserve edges, we propose an adaptive total variation term based on a weighted function, values of which are adapted to the features of pixels. The mean curvature is used to describe the features of images and adjusts values of the weighted function, i.e. strength of smoothing of the model. The Split Bregman is applied to solve the minimization problem. The numerical results demonstrate that, the proposed model yields better denosing performance compared with the classical model in both terms of quantitative and qualitative criteria.

Key words - image denoising; total variation; Split Bregman; ROF model

1. Đặt vấn đề

Trong quá trình thu nhận, truyền và ghi dữ liệu, hình ảnh sẽ bị nhiễu do khiếm khuyết trong cấu tạo của thiết bị thu phát hình ảnh hoặc do môi trường bên ngoài. Do đó, việc khử các hiệu ứng nhiễu là cần thiết nhằm phục vụ cho các bước xử lý ảnh tiếp theo.

Gọi $u: \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ là ảnh gốc (không nhiễu) và f là ảnh nhiễu của u . Quá trình gây nhiễu ảnh được mô hình hóa như sau:

$$f = u + \eta, \quad (1)$$

Trong đó, $\eta \sim \mathcal{N}(0, \sigma^2)$ là nhiễu cộng Gauss với giá trị trung bình $\mu = 0$ và phương sai σ^2 .

Việc khôi phục ảnh gốc u được xem như một bài toán ngược (*inverse problem*) với ràng buộc yếu. Phương pháp thông thường để giải quyết các bài toán ngược là bổ sung thành phần ổn định hóa (*regularization term*) vào hàm năng lượng (*energy function*) để ràng buộc bài toán có nghiệm duy nhất. Mô hình Rudin-Osher-Fatemi (ROF) [1] là một trong những nghiên cứu nổi tiếng theo hướng này, và được biểu diễn như sau:

$$\min_u \left(\frac{1}{2} \int_{\Omega} (u - f)^2 dx + \lambda \int_{\Omega} |\nabla u| dx \right), \quad (2)$$

với ∇ - toán tử gradient; λ là hệ số ổn định hóa.

Thành phần thứ nhất của hàm năng lượng (2) đo lường “độ trung thực” (*fidelity*) của ảnh khử nhiễu so với ảnh nhiễu f dưới dạng khoảng cách Euclid; còn thành phần thứ hai được gọi là tổng biến phân (*total variation*), đóng vai trò của thành phần ổn định hóa. Thành phần tổng biến phân đặc trưng cho độ mượt của ảnh khử nhiễu, đồng thời cho phép sự tồn tại của đường biên (*edges*).

Nhược điểm chính của mô hình ROF là hiệu ứng bậc thang (*staircase effect*) xuất hiện tại các cực trị và biên của ảnh. Hiệu ứng bậc thang đặc trưng bởi việc hình thành các vùng ảnh có độ sáng đồng nhất. Để giải quyết vấn đề này, các

tác giả đã đề xuất kết hợp tổng biến phân và Non-local means (NLM) [2, 3]. Trong đó, NLM [4] đóng vai trò xác định trọng số cho thành phần ổn định hóa dựa trên sự tương đồng giữa các điểm ảnh. Trong nghiên cứu [5], một mô hình thích nghi đã được đề xuất với thành phần ổn định hóa có khả năng tự điều chỉnh tương ứng với đặc trưng của điểm ảnh.

Trong bài báo này, nhóm tác giả nghiên cứu mô hình khử nhiễu dựa trên hướng tiếp cận biến phân (variational method). Đóng góp chính của nghiên cứu nằm ở đề xuất thành phần tổng biến phân thích nghi (adaptive total variation) nhằm cải thiện hiệu quả khử nhiễu của mô hình ROF. So với các mô hình tổng biến phân thích nghi đã có, nhóm tác giả đề xuất sử dụng độ cong trung bình của một mặt ảnh để mô tả đặc trưng của ảnh và từ đó điều chỉnh độ ổn định hóa (hay độ mịn ảnh) của mô hình thích nghi với tính chất của từng điểm ảnh.

2. Mô hình đề xuất

Nhóm tác giả đề xuất mô hình khử nhiễu sử dụng thành phần tổng biến phân thích nghi như sau:

$$\min_u \left(\frac{\lambda}{2} \int_{\Omega} (u - f)^2 dx + \int_{\Omega} \alpha(x) |\nabla u| dx \right), \quad (3)$$

với $\alpha(x)$ là hàm trọng số có giá trị biến đổi thích nghi với đặc trưng của từng điểm ảnh.

Mục tiêu của mô hình đề xuất là điều chỉnh độ làm mịn ảnh tùy thuộc vào đặc trưng của từng điểm ảnh. Cụ thể, đối với các vùng có độ sáng tương đồng, mô hình sẽ khử nhiễu mạnh, làm mịn ảnh; Còn đối với các vùng tại các đường biên, độ làm mịn ảnh sẽ giảm để giữ được các chi tiết ảnh. Để đạt được mục tiêu này, nhóm tác giả xây dựng hàm trọng số $\alpha(\cdot)$ dựa trên phép đo độ cong trung bình (mean curvature) như sau:

Một ảnh u xác định trên miền $\Omega \in \mathbb{R}^2$ có thể được xem như một mặt (*surface*) trong $\mathbb{R}^3$ được xác định bởi $z = f(x, y)$ với $(x, y) \in \Omega$. Độ cong trung bình [6] của một

mặt ảnh được định nghĩa như sau:

$$\kappa_u = \nabla \cdot \left(\frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} \right). \quad (4)$$

Độ cong trung bình của một mặt mô tả các đặc trưng hình học của mặt đó, bao gồm các góc, đường biên và độ tương phản [6]. Để giảm sự ảnh hưởng của nhiễu, nhóm tác giả áp dụng phép đo độ cong trung bình cho ảnh được lọc bởi bộ lọc Gauss:

$$\kappa_u = \nabla \cdot \left(\frac{\nabla G_\sigma * u}{\sqrt{1 + |\nabla G_\sigma * u|^2}} \right), \quad (5)$$

với, G_σ là bộ lọc Gaussian với tham số σ ; $*$ là phép chập.

Hàm trọng số được xác định bởi biểu thức:

$$\alpha(x) = \frac{1}{1 + \beta \bar{\kappa}_u}, \quad (6)$$

với, β là hệ số dương; $\bar{\kappa}_u$ là độ cong trung bình chuẩn hóa và được biểu diễn bởi:

$$\bar{\kappa}_u = \frac{\kappa_u}{\max(\kappa_u)}. \quad (7)$$

Do $\bar{\kappa}_u \in [0, 1]$ nên hàm $\alpha(x)$ đồng biến giảm từ 1 đến $1/(1 + \beta)$. Đối với các vùng có độ sáng tương đồng, $\bar{\kappa}_u \rightarrow 0$ nên $\alpha(x) \rightarrow 1$. Còn tại các đường biên, $\bar{\kappa}_u \rightarrow 1$ nên $\alpha(x) \rightarrow 1/(1 + \beta)$. Như vậy, hàm $\alpha(x)$ điều chỉnh trọng số của thành phần ổn định hóa trong hàm năng lượng (3), từ đó tăng độ làm mịn trong các vùng có độ sáng tương đồng và giảm độ làm mịn tại các đường biên.

3. Tính hội tụ

Để chứng minh tính hội tụ của nghiệm của bài toán (3), nhóm tác giả đưa ra định nghĩa của tổng biến phân thích nghi như sau.

Định nghĩa 1: Cho $\alpha(x) \geq 0$ là một hàm số thực trên miền Ω . Tổng biến phân thích nghi (TV_α) của một ảnh u trên miền Ω được định nghĩa bởi:

$$\begin{aligned} TV_\alpha &= \int_{\Omega} \alpha(x) |\nabla u| dx \\ &= \sup \left\{ \int_{\Omega} u \operatorname{div}(\xi) dx : \xi \in C_0^1(\Omega, B^2), |\xi_i(x)| \right. \\ &\quad \left. \leq \alpha(x), \forall x \in \Omega \right\}, \end{aligned} \quad (8)$$

với, $\sup$ là cận trên đúng (supremum); $C_0^1(\Omega, B^2)$ là không gian compact của các hàm khả vi liên tục trên miền Ω ; B^2 là hình tròn đơn vị mở trong $\mathbb{R}^2$; và div là toán tử phân kỳ (divergence).

Từ định nghĩa trên, ta chứng minh sự tồn tại và tính duy nhất của nghiệm của bài toán (3) như sau.

Định lý 1: Giả sử $u_n(x, y) \rightarrow u(x, y)$ trong không gian các hàm khả tích bậc 1 $L^1(\Omega)$, ta có:

$$\int_{\Omega} \lambda(x) |\nabla u| \leq \liminf_{n \rightarrow \infty} \int_{\Omega} \lambda(x) |\nabla u_n|. \quad (9)$$

Chứng minh. Theo định nghĩa của TV_α , đối với $\xi \in C_c^1(\Omega, B^2)$ bất kỳ, ta có:

$$\begin{aligned} \int_{\Omega} u \operatorname{div}(\xi) dx &= \lim_{n \rightarrow \infty} \int_{\Omega} u_n \operatorname{div}(\xi) dx \\ &\leq \lim_{n \rightarrow \infty} \int_{\Omega} \lambda(x) |\nabla u_n|. \end{aligned} \quad (10)$$

Bằng cách lấy cận trên đúng (supremum) cho 2 vế, ta chứng minh được Định lý 1. $\square$

Định lý 2: Giả sử ảnh $f \in L^2(\Omega)$, nghiệm u^* của bài toán (3) là duy nhất trong $BV(\Omega)$, với BV là không gian của các hàm có biến phân chặn (bounded variation).

Chứng minh. Rõ ràng rằng, hàm TV_α là hàm lồi. Do hàm (3) là tổng có trọng số giữa hàm lồi TV_α và hàm lồi nghiêm ngặt (thành phần thứ nhất của hàm (3)) nên hàm (3) là hàm lồi nghiêm ngặt trong $BV(\Omega)$. Tính duy nhất của nghiệm được suy ra trực tiếp từ tính lồi nghiêm ngặt của hàm (3).

4. Phương pháp tính

Nhóm tác giả áp dụng phương pháp Split Bregman [7] để giải bài toán tối ưu (3). Do tính rời rạc của ảnh số, nhóm tác giả lựa chọn dạng rời rạc của các toán tử vector như sau: Gọi $u_{i,j}$ ($i = 1, \dots, N$ và $j = 1, \dots, M$) là ảnh rời rạc. Toán tử gradient $\nabla u = ((\nabla u)_{i,j}^1, (\nabla u)_{i,j}^2)$ được xác định bởi:

$$\begin{aligned} (\nabla u)_{i,j}^1 &= \begin{cases} u_{i+1,j} - u_{i,j} & \text{nếu } i < N \\ 0 & \text{nếu } i = N \end{cases}, \\ (\nabla u)_{i,j}^2 &= \begin{cases} u_{i,j+1} - u_{i,j} & \text{nếu } j < M \\ 0 & \text{nếu } j = M \end{cases}. \end{aligned} \quad (11)$$

Toán tử phân kỳ được định nghĩa tương tự như dạng liên tục: $\operatorname{div} = -\nabla^*$, với ∇^* là đối ngẫu (adjoint) of ∇ . Ta có:

$$(\operatorname{div} p)_{i,j} = (\operatorname{div} p)_{i,j}^1 + (\operatorname{div} p)_{i,j}^2, \quad (12)$$

với

$$\begin{aligned} (\operatorname{div} p)_{i,j}^1 &= \begin{cases} p_{i,j}^1 - p_{i-1,j}^1 & \text{nếu } 1 < i < N \\ p_{i,j}^1 & \text{nếu } i = 1 \\ -p_{i-1,j}^1 & \text{nếu } i = N \end{cases}, \\ (\operatorname{div} p)_{i,j}^2 &= \begin{cases} p_{i,j}^2 - p_{i,j-1}^2 & \text{nếu } 1 < j < M \\ p_{i,j}^2 & \text{nếu } j = 1 \\ -p_{i,j-1}^2 & \text{nếu } j = M \end{cases}. \end{aligned} \quad (13)$$

Bài toán tối ưu (3) được biểu diễn lại dưới dạng bài toán có ràng buộc như sau:

$$\min_{d,u} \left(\alpha(x) |d| + \frac{\lambda}{2} \|u - f\|_2^2 \right) \text{ s.t. } d = \nabla u. \quad (14)$$

Bằng cách bổ sung thành phần ràng buộc, ta có bài toán không ràng buộc như sau:

$$\min_{d,u} \left(\alpha(x) |d| + \frac{\lambda}{2} \|u - f\|_2^2 + \frac{\mu}{2} \|d - \nabla u\|_2^2 \right). \quad (15)$$

Áp dụng vòng lặp Bregman, ta có:

$$\begin{aligned} \min_{d,u} \left(\alpha(x) |d| + \frac{\lambda}{2} \|u - f\|_2^2 \right. \\ \left. + \frac{\mu}{2} \|d - \nabla u - b\|_2^2 \right), \end{aligned} \quad (16)$$

với, b là một tham số được cập nhật sau mỗi vòng lặp Bregman; μ – hệ số dương.

Ta giải bài toán (16) bằng cách lần lượt tối ưu bài toán

theo từng biến u và d riêng lẻ đến khi đạt điều kiện hội tụ. Như vậy, ta có 2 bài toán con cần được giải quyết như sau:

- Bài toán con theo d : Bằng cách cố định biến u , ta có bài toán theo biến d như sau:

$$d^{k+1} = \min_d \left(\alpha(x)|d| + \frac{\mu}{2} \|d - \nabla u^k - b^k\|_2^2 \right). \quad (17)$$

Giá trị tối ưu của d được xác định bằng toán tử *shrink* [8]:

$$d^{k+1} = \text{shrink} \left(\nabla u^k + b^k, \frac{\alpha(x)}{\mu} \right), \quad (18)$$

với

$$\text{shrink}(x, \gamma) = \frac{x}{|x|} * \max(|x| - \gamma, 0). \quad (19)$$

- Bài toán con theo u : Bằng cách cố định biến d , ta có bài toán theo biến u như sau:

$$u^{k+1} = \min_u \frac{\lambda}{2} \|u - f\|_2^2 + \frac{\mu}{2} \|d^{k+1} - \nabla u - b^k\|_2^2, \quad (20)$$

Điều kiện đạt cực trị của bài toán trên được xác định bởi phương trình:

$$(\lambda I - \mu \Delta)u^{k+1} = \lambda f + \mu \nabla^*(d^{k+1} - b^k), \quad (21)$$

với $\text{div} := -\nabla^*$ và $\Delta := \text{div} \nabla$; I – ma trận đơn vị.

Phương trình (21) có thể được giải trong miền Fourier như sau:

$$u^{k+1} = \mathcal{F}^{-1} \left(\frac{\mathcal{F}(\lambda f + \mu \nabla^*(d^{k+1} - b^k))}{\lambda I - \mu \mathcal{F}(\Delta)} \right), \quad (22)$$

với $\mathcal{F}$ and $\mathcal{F}^{-1}$ lần lượt là biến đổi Fourier thuận và ngược.

Biến b được cập nhật sau mỗi vòng lặp Bregman như sau:

$$b^{k+1} = b^k + \nabla u^{k+1} - d^{k+1}. \quad (23)$$

Thuật toán được tổng hợp lại như sau:

Thuật toán Split Bregman để giải bài toán (16)

- 1: **Khai báo:** $u^0 = f, d^0 = \mathbf{0}, b^0 = \mathbf{0}$
- 2: **While** “chưa thỏa điều kiện hội tụ” **do**
- 3: Tính d^{k+1} theo (18)
- 4: Tính u^{k+1} theo (22)
- 5: Tính b^{k+1} theo (23)
- 6: **end**

5. Kết quả mô phỏng và đánh giá

Trong mục này, nhóm tác giả so sánh mô hình đề xuất (viết tắt là ATVD) với mô hình ROF [1]. Nhóm tác giả đánh giá hiệu quả của các mô hình dựa trên 3 tiêu chí: PSNR, SSIM và đánh giá trực quan. Mỗi tiêu chí đo lường một khía cạnh khác nhau của kết quả khử nhiễu. Trong khi PSNR đánh giá sự chênh lệch cường độ sáng giữa ảnh khử nhiễu và ảnh gốc thì SSIM đo lường sự tương đồng về cấu trúc ảnh. Gọi f và u là ảnh không nhiễu và ảnh khử nhiễu. Các đại lượng PSNR và SSIM được định nghĩa như sau:

$$PSNR = 10 \log_{10} \left(\frac{I_{max}^2}{\frac{1}{MN} \|u - f\|_2^2} \right), \quad (24)$$

$$SSIM = \frac{(2\mu_f\mu_u + c_1)(2\sigma_{f,u} + c_2)}{(\mu_f^2 + \mu_u^2 + c_1)(\sigma_f^2 + \sigma_u^2 + c_2)},$$

với I_{max} là giá trị cường độ sáng lớn nhất của ảnh nhiễu f ; M và N là kích thước của ảnh; μ_f, μ_u, σ_f và σ_u lần lượt là giá trị trung bình và độ lệch chuẩn của f và u ; c_1 và c_2 là các hằng số.

Các mô hình được đánh giá trên 4 ảnh thông dụng (Hình 1). Các ảnh được làm nhiễu bởi nhiễu cộng Gauss với giá trị trung bình bằng 0 và độ lệch chuẩn σ . Ba mức độ nhiễu $\sigma = 10, 20, 30$ được xem xét.



Hình 1. Ảnh kiểm tra. Từ trái qua phải: Lena, Cameraman, Flinstones, Peppers

Bảng 1 trình bày kết quả PSNR và SSIM của 2 mô hình. Kết quả tốt nhất cho từng ảnh được tô đậm. Bảng 1 cho thấy ATVD cho kết quả PSNR and SSIM tăng lên trung bình khoảng 0,34dB và 0,0042. Kết quả được cải thiện nhiều hơn là đối với các ảnh “Cameraman” và “Flinstones”. Điều này có thể lý giải là do các ảnh “Cameraman” và “Flinstones” đặc trưng bởi các vùng có độ sáng tương đồng với kích thước lớn, phù hợp với hoạt động của hàm trọng số trong mô hình ATVD. Còn đối với ảnh “Lena”, mức độ cải thiện về chỉ số PSNR và SSIM không đáng kể.

Bảng 1. Kết quả PSNR và SSIM của các mô hình (kết quả được trình bày với định dạng PSNR(SSIM))

	Lena	Camera man	Flinstones	Peppers
$\sigma = 10$				
ROF	33,32 (0,8843)	30,61 (0,8737)	29,44 (0,8758)	32,25 (0,9097)
ATVD	33,55 (0,8866)	31,13 (0,8786)	29,87 (0,8799)	32,64 (0,9122)
$\sigma = 20$				
ROF	30,75 (0,8361)	27,55 (0,8095)	25,75 (0,8100)	28,89 (0,8560)
ATVD	30,84 (0,8371)	27,91 (0,8139)	26,20 (0,8183)	29,17 (0,8587)
$\sigma = 30$				
ROF	29,33 (0,8037)	25,86 (0,7654)	23,86 (0,7621)	27,14 (0,8174)
ATVD	29,43 (0,8001)	26,30 (0,7701)	24,25 (0,7722)	27,52 (0,8189)

Hình 2 mô tả kết quả khử nhiễu của ảnh “Lena” đối với mức độ nhiễu $\sigma = 20$. Ảnh phóng to (hàng thứ ba) cho thấy, ATVD cho kết quả khử nhiễu tốt hơn với vùng mặt mịn hơn và ít tạo ra hiệu ứng bậc thang hơn so với mô hình ROF. Điều này được thể hiện rõ hơn ở các ảnh đường đồng mức (contour lines). Có thể thấy rằng, ATVD làm giảm hiệu ứng bậc thang tại các vùng má và trán. Tóm lại, ATVD cải thiện kết quả khử nhiễu của mô hình ROF trên cả hai tiêu chí định lượng và định tính.



(a) Ảnh gốc



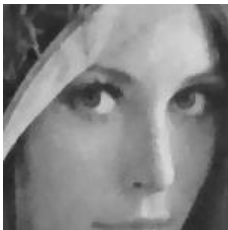
(b) Ảnh nhiễu



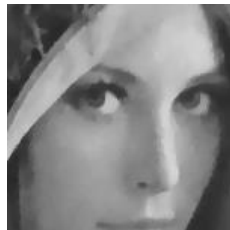
(c) ROF



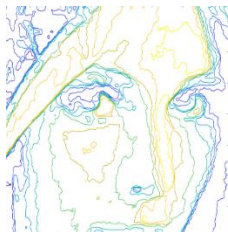
(d) ATVD



(e) ROF



(f) ATVD



(g) ROF



(h) ATVD

6. Kết luận

Trong bài báo này, chúng tôi đã đề xuất mô hình khử nhiễu ảnh sử dụng tổng biến phân thích nghi. Bằng cách sử dụng hàm trọng số dựa trên phép đo độ cong trung bình, mô hình đề xuất có khả năng điều chỉnh độ làm mịn ảnh tùy thuộc vào đặc trưng của từng điểm ảnh. Kết quả mô phỏng cho thấy mô hình đề xuất cho kết quả khử nhiễu tốt hơn so với mô hình ROF theo cả tiêu chí định lượng và định tính.

TÀI LIỆU THAM KHẢO

- [1] Rudin, Leonid I., Stanley Osher, and Emad Fatemi. "Nonlinear total variation based noise removal algorithms". *Physica D: nonlinear phenomena* 60.1-4 (1992): 259-268.
- [2] Liu, Xinwu, and Lihong Huang. "A new nonlocal total variation regularization algorithm for image denoising". *Mathematics and Computers in Simulation* 97 (2014): 224-233.
- [3] Hu, Haijuan, and Jacques Froment. "Nonlocal total variation for image denoising". *2012 Symposium on Photonics and Optoelectronics*. IEEE, 2012.
- [4] Buades, Antoni, Bartomeu Coll, and J-M. Morel. "A non-local algorithm for image denoising". *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. Vol. 2. IEEE, 2005.
- [5] Zhou, Weifeng, and Qingguo Li. "Adaptive total variation regularization based scheme for Poisson noise removal". *Mathematical Methods in the Applied Sciences* 36.3 (2013): 290-299.
- [6] Zhu, Wei, and Tony Chan. "Image denoising using mean curvature of image surface". *SIAM Journal on Imaging Sciences* 5.1 (2012): 1-32.
- [7] Chen, Y., and T. Wunderli. "Adaptive total variation for image restoration in BV space". *Journal of Mathematical Analysis and Applications* 272.1 (2002): 117-137.
- [8] Setzer, Simon. "Operator splittings, Bregman methods and frame shrinkage in image processing". *International Journal of Computer Vision* 92.3 (2011): 265-280.

Hình 2. Ảnh khử nhiễu của ảnh "Lena" đối với mức độ nhiễu $\sigma = 20$. Hàng thứ hai: ảnh kích thước đầy đủ; hàng thứ ba: ảnh phóng to; hàng thứ tư: ảnh đường đồng mức (contour lines) tương ứng với các ảnh phóng to

(BBT nhận bài: 20/01/2020, hoàn tất thủ tục phản biện: 10/3/2020)